

Security analysis of the decoy method with the Bennett-Brassard 1984 protocol for finite key lengths

Masahito Hayashi^{1,2} and Ryota Nakayama^{3,1}

¹ Graduate School of Mathematics, Nagoya University, Furocho, Chikusa-ku, Nagoya, 464-860 Japan

² Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117542

³ Graduate School of Information Sciences, Tohoku University, Aoba-ku, Sendai, 980-8579, Japan

Abstract. This paper provides the sacrifice bit length with the Bennett-Brassard 1984 protocol for finite key lengths when we employ the decoy method. Using the method, we can guarantee the security parameter for realizable quantum key distribution system. The generated key rates with finite key lengths are numerically evaluated. The proposed method improves the key generation rate even in the asymptotic setting. Further, it can perfectly estimate the additional channel parameters for decoy method in the asymptotic limit when the decoy intensity is infinitesimal.

PACS numbers: 03.67.Dd, 03.67.Hk, 03.67.-a, 05.30.Jp

Keywords: decoy method, finite-key length, BB84 protocol, phase error, interval estimation, percent point

1. Introduction

Quantum key distribution (QKD) protocol proposed by Bennett-Brassard [1] is one of the most applicable protocols in quantum information. In the ideal setting, this protocol is trivially secure. However, in the real setting, there are two obstacles for security. One is noise of the communication quantum channel. Due to the presence of the noise, the eavesdropper can leak a part of information behind the noise. The second one is the imperfection of the photon source. If the sender sends the two photon state instead of the single photon state, the eavesdropper can leak one photon so that she can leak information perfectly. Indeed, many realized QKD has been realized with weak coherent source. In this case, the photon number of generated state obeys the Poisson distribution of average μ , which is called the intensity. The first problem can be resolved by combination of the error correction and random privacy amplification [2, 3, 4, 5]. Shor-Preskill [2] and Mayes [3] showed that this method gives the secure keys asymptotically. Gottesman-Lo-Lütkenhaus-Preskill (GLLP)[6] extended their result to the case when the photon source has the imperfection. However, GLLP's result assumes the breakdown of photon numbers among received quantum states. Indeed, there is possibility that the eavesdropper can control the receiver's detection rate depending on the photon number because the different photon number states can be distinguished by the eavesdropper. In order to resolve this problem, we need to estimate the detection rate among the single photon states. Hwang proposed the decoy method to estimate it, in which, the sender randomly changes the intensity [7]. This method has been improved by many researchers[8, 9, 10, 11, 12, 13, 14]. In this method, in order to estimate the detection rates, the sender randomly chooses several kinds of states with different intensities. One is the signal states, from which, we generate the secret keys. Others are the decoy states, which are used for estimating the action of eavesdropper and have different intensity from the signal state.

However, we cannot still realize a truly secure QKD system in the real world due to the finiteness of the realized code length. Most of the above results assume the asymptotic setting except for Mayers[3]. Also, their privacy amplification requires many calculation times. Renner [15] proposed to use universal₂ hash function for privacy amplification and showed the security under this kind of hash function. Universal₂ hash function has been recognized as a fundamental tool for information theoretical security [16, 17, 18, 19]. His security proof is quite different from the traditional Shor-Preskill formalism because it employs the left over hashing lemma (privacy amplification) while the traditional Shor-Preskill formalism employs error correction. He proved security based on the On the other hand, in the context of the traditional Shor-Preskill formalism, it was shown that the leaked information can be evaluated only by the phase error probability[5, 20, 21, 22, 23], which implies that the phase error correction guarantees the security. Using this fact, Tsurumaru-Hayashi [24] showed that the security under a wider class of hash function, which is called ε -almost dual universal₂ hash function.

In order to treat the finiteness problem, In the single photon case, Hayashi treated

the second order analysis of the coding rate with Gaussian approximation by using the above phase error correction formalism[5]. Indeed, recently, the second order analysis has been paid much attentions to among information theory community due to the relation with analysis of finite-length code[25, 26, 27, 28]. Scarani et al.[29] and Sano et al. [30] also treated this problem only for collective attack. Recently, using Renner's formalism, Tomamichel et al [31] derived an upper bound formula for security with the finite-length code. Using the phase error correction formalism, Hayashi-Tsurumaru [32] derived better upper bound formulas for security with the finite-length code, which attain the second order rate given in Hayashi [5]. They also treated the variable length type of the universal composability based on the phase error correction formalism. However, these results assume the single photon source. Furrer et al.[33] gave a finite-length analysis with continuous variable quantum key distribution, which works with weak coherent source.

While continuous variable quantum key distribution can be implemented with a cheap Homodyne detection, the decoy method with BB84 protocol can achieve the longest distance with the current technology[34, 35]. Hence, employing the phase error correction formalism, we treat the security of finite-length code of BB84 protocol when we use the weak coherent light and the decoy method. Our finite-length evaluation can be divided into two parts. The first part is the estimation of the channel parameters and the source parameters. Indeed, all of the eavesdropper's action can be described by the channel parameters, which contains the detection rate and error rate for each number state. Hence, it is sufficient for the evaluation of the leaked information to estimate the channel parameters. Since the photon number of the source has a stochastic behavior, we have to treat the stochastic behavior of photon numbers in sources as well as that of the estimator of the channel parameter. The second part is the evaluation of the virtual decoding phase error probability under the assumption that the above parameters belong to a certain region. Combining two parts, depending on the intensities of signal and decoy pulses, we derive a sacrifice bit-length that guarantees that the universal composability of the final keys is less than a certain threshold. As is illustrated in Fig. 3, our calculation formula for a sacrifice bit-length employs only the basic formulas of the *percentage points* and the *interval estimation* of the binomial distribution, whose packages are available in several computer systems. Hence, it does not contain any optimization process, and then it requires a relatively smaller calculation time. Further, we numerically calculate the key generation rate per pulses in several cases. In these numerical calculations, since the required error probabilities are too small to calculate the exact percentage points and the exact interval estimation, we employ Chernoff bound, which is summarized in Appendix. Since the required error probabilities are too small, the difference between the exact value and the value based on Chernoff bound is sufficiently small. We also improve the asymptotic key generation rate when we mix the vacuum pulse and one decoy intensity pulse to the signal pulse.

Our method has improvement even in the asymptotic setting because our parameterization is different from existing parameters in the following way. Existing

methods estimate the counting rates (yields) of the respective states (the vacuum state, the single photon state, etc.) and the probabilities that the photon is detected and the phase error occurs for the respective states as channel parameters. However, our method estimate the two kinds of probabilities, the probabilities that the photon is detected and the phase error occurs, and the probabilities that the photon is detected and the phase error does not occur, in the respective states. Our parametrization has the one-to-one correspondence with existing parametrization. However, as is shown in Subsection 8.2, our parametrization yields a better asymptotic key generation rate than existing parametrization. Usually, the decoy method estimates these parameters by using the constraint that these parameters are non-negative. The reason why our parameterization is better is that the non-negativity in our parameterization yields a more strict condition than that in existing parameterization. Hence, our parameterization provides a better key generation rate even in the asymptotic case. Further, in Subsection 8.2, we show that when the decoy intensity is infinitesimal, we can perfectly estimate these parameters for the single photon state in the asymptotic limit, i.e., when the length M of raw keys goes to infinity. Note that it is usually not impossible to perfectly estimate these parameters for the single photon state in the asymptotic limit while it is possible to perfectly estimate these parameters for the vacuum state in the asymptotic limit by using the vacuum decoy state. However, the convergence to the asymptotic key generation rate is not uniform when the decoy intensity is infinitesimal. That is, the convergence speed depends on the decoy intensity. Hence, it is needed to address the trade-off between the speed of the convergence and the asymptotic key generation rate. This trade-off is discussed by considering the second order asymptotics in Section 11. In Subsection 13.1, we numerically check how well the approximation by the second order asymptotics works.

This paper also addresses the case when the intensities are not fixed or are different from our intent. We treat the asymptotic key generation rate when the intensities are different from our intent in Subsection 8.3. We numerically demonstrate how error of our guess influences the asymptotic key generation error by considering the error rate of our guess of the intensities. In Subsection 12.1, we discuss our key generation rate with finite-length when the intensities are not fixed and obey certain probability distributions. In Subsection 13.2, We numerically calculate the above key generation rate when the intensities obey Gaussian distributions. In Subsection 12.2, we discuss our key generation rate with finite-length when the intensities are different from our intent. More precisely, we consider the case when there are several candidates for distributions of the intensities.

The organization of the remaining part is the following. As a preparation, Section 2 reviews the result for the universal composability of the final keys when we know the breakdown of the received pulses and the phase error probability among single photon pulses. Section 3 gives a concrete protocol of the decoy method. Section 4 explains how eavesdropper's action can be described. Section 5 summarizes a fundamental knowledge for random variables. Section 6 briefly describes our security proof the outline

of the latter discussion. Section 7 gives the estimate of channel parameters when the breakdown of the generated sources is given. In Section 8, we apply the method given in Section 7 to the asymptotic case. In Subsection 8.2, we compare our asymptotic key generation rate with the existing rate. In Subsection 8.3, similar to [13, 14], we treat the asymptotic key generation rate when we cannot identify the true intensity perfectly. In Section 9, we treat stochastic fluctuation of the photon number of the sources. Then, combining the discussions in Sections 6, 7, and 9, we obtain the sacrifice bit-length guaranteeing the security. However, the obtained sacrifice bit-length can be improved by the way given in Section 10. In the improved method, we put out several probabilities from the square root. Since the improved method is too complicated, we give a looser evaluation, first. After describing whole structure, we give a better method. Indeed, when the decoy intensity is less than the signal intensity, the asymptotic key generation rate is best when the decoy intensity goes to zero. However, in the finite-length setting, the estimation does not work properly under the limit. In Section 11, we take the asymptotic expansion of sacrifice bit length up to the second order. Then, we optimize the decoy intensity in the sense of the second order coefficient. In Section 12, we treat the finite sacrifice bit-length when the source intensity is not fixed. In Section 13, we give the several numerical calculations with the finite-length setting. In Appendices A,B, and C, we summarize the basic knowledge concerning the tail probability and the interval estimation under the binary distribution. In Appendix D, we summarize calculations required for the numerical calculation in Section 13.

2. Security evaluation

An evaluation method to use the trace norm is known as a security criterion in QKD, which is often called the universal composability[36]. When the length m of the final keys is not fixed, we need a more careful treatment. We denote the final state and Eve's final state by $\rho_{AE|m}$ and $\rho_{E|m}$, respectively when the length of the final keys is m . Our ideal Alice's state is the uniform distribution $\rho_{\text{mix}|m}$ on m bits.

Hence, the ideal composite state is $\rho_{\text{mix}|m} \otimes \rho_{E|m}$. We denote the state indicating that the length of final keys is m , by $|m\rangle\langle m|$, and its probability by $P(m)$. Then, the state of the composite system is $\rho_{AE} := \sum_m P(m) |m\rangle\langle m| \otimes \rho_{AE|m}$, and its ideal state is $\rho_{\text{ideal}} := \sum_m P(m) |m\rangle\langle m| \otimes \rho_{\text{mix}|m} \otimes \rho_{E|m}$. Hence, the averaged universal composability of the obtained keys is written as the trace norm of the difference between the real state ρ_{AE} of the composite system and its ideal state ρ_{ideal} [37]D

$$\|\rho_{A,E} - \rho_{\text{ideal}}\|_1 \quad (1)$$

Hence, a smaller trace norm guarantees more secure final keys.

On the other hand, when we apply surjective hash functions as the privacy amplification [21], [32, (10)] the above value is bounded by the averaged virtual phase error probability P_{ph} as.

$$\|\rho_{A,E} - \rho_{\text{ideal}}\|_1 \leq 2\sqrt{2}\sqrt{P_{ph}} \quad (2)$$

Then, the security analysis of QKD can be reduced to the evaluation of P_{ph} .

In the following, we consider the protocol with the privacy amplification with the sacrifice bit length S over the raw key with length M . When the number of phase error among M bits-raw keys is E and we apply the minimum length decoding, the averaged virtual phase error probability P_{ph} is evaluated as

$$P_{ph} \leq 2^{Mh(\min(\frac{E}{M}, \frac{1}{2})) - S}. \quad (3)$$

Hence, we can guarantee the security of the final keys when the sacrifice bit length S is sufficiently larger than $Mh(\min(\frac{E}{M}, \frac{1}{2}))$. However, the number E of errors among raw keys does not take a deterministic value, it obeys a probability distribution $Q(E)$. Then, when we apply the minimum length decoding, the averaged virtual phase error probability P_{ph} is evaluated as

$$E(P_{ph}) \leq \sum_E Q(E) \min(2^{Mh(\min(\frac{E}{M}, \frac{1}{2})) - S}, 1). \quad (4)$$

Next, we consider the case when the transmitted pulses generating the raw keys of M bits take the following three types. The first is the vacuum state, the second is the single-photon state, and the third one is the multi-photon state. In the following, we assume that the transmitted pulses generating the raw keys of M bits consist of J^0 bits pulses with the vacuum state, J^1 bits pulses with the single-photon state, and J^2 bits pulses with the multi-photon state. Due to the assumption, the relation $M = J^0 + J^1 + J^2$ holds.

When we send the pulse with the vacuum state, no information can be leaked to Eve. That is, the leaked information to Eve equals to the leaked information to Eve when we send the pulse with the single-photon state and with phase error probability 0. On the other hand, in the case of the multi-photon case, we have to consider that all information is leaked to Eve. Hence, the leaked information to Eve equals to the leaked information to Eve when we send the pulse with the single-photon state and with phase error probability $1/2$. Therefore, when we denote the number of phase errors among J^1 bits, after we apply a proper class of hash functions in the privacy amplification[‡], the averaged virtual phase error probability P_{ph} is evaluated as [21, (19)] and [24]§

$$P_{ph} \leq 2^{\phi(J^0, J^1, J^3) - S} \quad (5)$$

because $J^2 = M - J^0 - J^1$, where we define

$$\phi(J^0, J^1, J^3) := J^1 h(\min(\frac{J^3}{J^1}, \frac{1}{2})) + (M - J^0 - J^1). \quad (6)$$

Due to (5), we can regard $\phi(J^0, J^1, J^3)$ as a leaked information.

[‡] More precisely, when we apply ε -almost dual universal₂ hash functions, P_{ph} is evaluated as $P_{ph} \leq \varepsilon \cdot 2^{\phi(J^0, J^1, J^3) - S}$. As is explained in [24], several practical hash functions, e.g., the concatenation of Toeplitz matrix and the identity matrix, are 1-almost dual universal₂.

[§] In the derivation[21, (19)], we consider that The J^1 qubits has the phase error rate $\min(\frac{J^3}{J^1}, \frac{1}{2})$ and The $J^2 (= M - J^0 - J^1)$ qubits has the phase error rate $1/2$.

In the real case, the values J^0 , J^1 , and J^3 do not take deterministic values, and obey the probability distribution $Q(J^0, J^1, J^3)$. Hence, the averaged virtual phase error probability P_{ph} is evaluated by

$$P_{ph} \leq \sum_{J^0, J^1, J^3} Q(J^0, J^1, J^3) \min(2^{\phi(J^0, J^1, J^3) - S}, 1). \quad (7)$$

In the general case, the size of sacrifice bit length S also does not take a deterministic value, and is stochastically determined. In such a case, the values J^0 , J^1 , J^3 , and S obey a simultaneous distribution $Q(J^0, J^1, J^3, S)$, and the averaged virtual phase error probability P_{ph} is evaluated by

$$P_{ph} \leq \sum_{J^0, J^1, J^3, S} Q(J^0, J^1, J^3, S) \min(2^{\phi(J^0, J^1, J^3) - S}, 1). \quad (8)$$

In the following, for a simplicity, we employ the notations $\mathbf{J} = (J^0, J^1, J^3)$ and $\phi(\mathbf{J}) := \phi(J^0, J^1, J^3)$.

3. Decoy method

In the following, we assume that the raw keys with M bits are generated by N bits pulse transmission with imperfect photon source. Now, we assume that there are N^0 pulses with the vacuum state and N^1 pulses with the single-photon state among N transmitted pulses

Then, using the detection probability \bar{q}^0 of the vacuum state transmission, the detection probability \bar{q}^1 of the single-photon state transmission, and the probability \bar{b}_\times^1 when we detect the pulse and the phase error occurs for the single-photon state transmission, the numbers J^0 , J^1 , and J^3 can be estimated as

$$J^0 \sim N^0 \bar{q}^0, \quad J^1 \sim N^1 \bar{q}^1, \quad J^3 \sim N^1 \bar{b}_\times^1. \quad (9)$$

However, it is not easy to estimate the breakdown. Now, we consider the case when the N weak coherent pulses with intensity μ_1 are transmitted. Then, we obtain the expansion concerning the photon-number states.

$$\sum_{n=0}^{\infty} e^{-\mu_1} \frac{\mu_1^n}{n!} |n\rangle \langle n| = e^{-\mu_1} |0\rangle \langle 0| + e^{-\mu_1} \mu_1 |1\rangle \langle 1| + e^{-\mu_1} \mu_1^2 \omega_2 \rho_2, \quad (10)$$

where

$$\rho_2 := \frac{1}{\omega_2} \sum_{n=2}^{\infty} \frac{\mu_1^{n-2}}{n!} |n\rangle \langle n|, \quad \omega_2 := \frac{1}{\mu_1^2} (e^{\mu_1} - (1 + \mu_1)). \quad (11)$$

Then, the breakdown can be estimated as

$$N^0 \sim N e^{-\mu_1}, \quad N^1 \sim N e^{-\mu_1} \mu_1. \quad (12)$$

Hence, it is needed to estimate the parameters \bar{q}^0 , \bar{q}^1 , and \bar{b}_\times^1 . For this purpose, we shuffle coherent pulses with another intensity μ_2 . In the following, we assume that

$\mu_1 < \mu_2$. This method is called the decoy method [7, 8, 9, 10, 11]||D Indeed, the coherent pulse with another intensity μ_2 has the following expansion.

$$\begin{aligned} \sum_{n=0}^{\infty} e^{-\mu_2} \frac{\mu_2^n}{n!} |n\rangle \langle n| &= e^{-\mu_2} |0\rangle \langle 0| + e^{-\mu_2} \mu_2 |1\rangle \langle 1| \\ &+ e^{-\mu_2} \mu_2^2 \omega_2 \rho_2 + e^{-\mu_2} \mu_2^2 (\mu_2 - \mu_1) \omega_3 \rho_3, \end{aligned} \quad (13)$$

where

$$\begin{aligned} \rho_3 &:= \frac{1}{\omega_3} \sum_{n=3}^{\infty} \frac{\mu_2^{n-2} - \mu_1^{n-2}}{(\mu_2 - \mu_1) n!} |n\rangle \langle n| \\ \omega_3 &:= \frac{1}{\mu_2^2} (e^{\mu_2} - (1 + \mu_2 + \frac{\mu_2^2}{2})) - \frac{1}{\mu_1^2} (e^{\mu_1} - (1 + \mu_1 + \frac{\mu_1^2}{2})). \end{aligned} \quad (14)$$

Using the difference between the two expansion coefficients, we can estimate the detection rates \bar{q}^0 and \bar{q}^1 .

In the following, we give a detail of our protocol.

- (1):**Transmission** Alice (the sender) sends the pulses with the vacuum, the coherent pulses with the intensity μ_1 , and the coherent pulses with the intensity μ_2 , randomly with a certain rate. Here, we choose the bit basis and the phase basis with the ratio $1 - \lambda : \lambda$ among the coherent pulses with the intensity μ_1 , and the coherent pulses with the intensity μ_2 ,
- (2):**Detection** Bob (the receiver) chooses the bit basis and the phase basis with the ratio $1 - \lambda : \lambda$ and measures the detected pulses. Then, he records existence or non-existence of the detection, his basis, and the measured bit.
- (3):**Verification of basis** Alice sends Bob all information concerning the basis and the intensity for all pulses. Bob sends Alice what pulses has the coincidence basis. Then, we denote the number of vacuum pulses, the number of pulses with the intensity μ_1 and the phase basis in the both sides, and the number of pulses with the intensity μ_2 and the phase basis in the both sides, by N_0 , N_1 , and N_2 , respectively. We also denote the number of pulses with the intensity μ_1 and the bit basis in the both sides and the number of pulses with the intensity μ_2 and the bit basis in the both sides, by N and N' , respectively. See Fig. 1.
- (4):**Parameter estimation** Alice and Bob announce all bit information concerning $N_1 + N_2$ pulses with the phase basis in the both sides. Then, we denote the number of vacuum pulses detected by Bob by M_0 . We also denote the number of coherent pulses detected by Bob with the intensity μ_i , the phase basis in the both sides, and the agreement bit (the disagreement bit) by $M_i(M_{i+2})$ for $i = 1, 2$. (However, we will not use M_4 .) Further, we denote the number of coherent pulses detected by Bob with the intensity μ_1 and the bit basis in the both sides and the number of coherent pulses detected by Bob with the intensity μ_1 and the bit basis in the both sides by M and M' , respectively. See Fig.2.

|| In a wider sense, we can regard the check bits estimating the phase error probability as another kind of decoy state.

In the following, we describe the key distillation protocol for M bits raw keys generated by the coherent pulses with intensity μ_1 . The key distillation protocol for M' bits raw keys generated by the coherent pulses with intensity μ_2 can be obtained N and M by N' and M' , respectively.

- (5):Error correction** First, we choose a suitable M -bits classical code C_1 that works for the expected bit error rate p_+ . We prepare a set $\{\mathbf{s}_{[s]}^2\}_{[s] \in \mathbb{F}_2^M/C_1}$ of representatives for decoding. We also prepare another a set $\{\mathbf{s}_{[s]}^1\}_{[s] \in \mathbb{F}_2^M/C_1}$ of representatives. Then, Alice and Bob exchange their information \mathbb{F}_2^M/C_2^\perp . Alice obtains $\mathbf{x} := \mathbf{s} - \mathbf{s}_{[s]}^1$ in C_2^\perp , and Bob obtains $\mathbf{x}' := \mathbf{s}' - \mathbf{s}_{[s]}^1 - \mathbf{s}_{[s'-s]}^2$ in C_1 .
- (6):Privacy amplification** By using the method explained latter, we define the sacrifice bit length S in the privacy amplification from $N, N_0, N_1, N_2, M, M_0, M_1, M_2, M_3$. Then, Alice and Bob apply ε -almost dual universal₂ hash function from $C_1 \cong \mathbb{F}_2^l$ to \mathbb{F}_2^{l-S} [24]. Then, they obtain the final keys.
- (7):Error verification** Alice and Bob apply a suitable hash function to the final keys. They exchange the exclusive OR between the above hash value and other prepared secret keys. If the above exclusive OR agrees, their keys agree with a high probability[38, 39].

Basis in Alice	Basis in Bob	Vacuum	μ_1	μ_2
Bit basis	Bit basis	N_0	N	N'
	Phase basis			
Phase basis	Phase basis		N_1	N_2
	Bit basis			

Figure 1. Breakdown of transmitted pluses

Basis in Alice	Basis in Bob		Vacuum	μ_1	μ_2
Bit basis	Bit basis		M_0	M	M'
	Phase basis				
Phase basis	Phase basis	correct		M_1	M_2
		incorrect		M_3	
	Bit basis				

Figure 2. Breakdown of detected pluses

In the error correction, we lose more than $Mh(p_+)$ bits. When we lose $\eta Mh(p_+)$ bits in the error correction, the final key length is $M - \eta Mh(p_+) - S$. Generally, we chose η to be 1.1.

4. Description of Eve

In the following, we describe the strategy of Eve. For this purpose, we treat only the vacuum pulses and the pulses with coincidence basis, whose total number of pulses is $N_0 + N_1 + N_2 + N + N'$. We do not treat other types of pulses. In this case, Eve cannot distinguish the intensities μ_1 and μ_2 perfectly. Alternatively, we assume that Eve can choose her strategy depending on the number of photons because she can distinguish the number of photons. Then, Eve is assumed to distinguish the states $|0\rangle\langle 0|$, $|1\rangle\langle 1|$, ρ_2 , and ρ_3 .

In the following, we assume the following:

- There are N_1^0 pulses with the vacuum state and N_1^1 pulses with the single-photon state among N_1 pulses with the intensity μ_1 and the phase basis.
- There are N_2^0 pulses with the vacuum state, N_2^1 pulses with the single-photon state, and N_2^2 pulses with the state ρ_2 among N_2 pulses with the intensity μ_2 and the phase basis.
- There are N^0 pulses with the vacuum state and N^1 pulses with the single-photon state among N pulses with the intensity μ_1 and the bit basis.
- There are $N^{0'}$ pulses with the vacuum state, $N^{1'}$ pulses with the single-photon state, and $N^{2'}$ pulses with the state ρ_2 among N' pulses with the intensity μ_2 and the bit basis.

For a simplicity, we employ the notations $\mathbf{N} := (N^0, N^1)$, $\mathbf{N}' := (N^{0'}, N^{1'}, N^{2'})$, $\mathbf{N}_1 := (N_1^0, N_1^1)$, $\mathbf{N}_2 := (N_2^0, N_2^1, N_2^2)$, and $\vec{\mathbf{N}} := (\mathbf{N}, \mathbf{N}', \mathbf{N}_1, \mathbf{N}_2)$.

In the above breakdown, there are $N_0 + N_1^0 + N_2^0 + N^0 + N^{0'}$ pulses with the vacuum state, $N_1^1 + N_2^1 + N^1 + N^{1'}$ pulses with the single-photon state, $N_1^2 + N_2^2$ pulses with the state ρ_2 and the phase basis, and N_2^3 pulses with the state ρ_3 and the phase basis, where $N_1^2 := N_1 - N_1^0 - N_1^1$ and $N_2^3 := N_2 - N_2^0 - N_2^1 - N_2^2$. Note that the average state with the bit basis is not the same as the average state with the phase basis in the case of the multi-photon state.

Then, Eve is assumed to control the detection rates in the Bob's side \bar{q}^0 , \bar{q}^1 , \bar{q}_\times^2 , and \bar{q}_\times^3 among $N_0 + N_1^0 + N_2^0 + N^0 + N^{0'}$ vacuum pulses, $N_1^1 + N_2^1 + N^1 + N^{1'}$ single-photon pulses, $N_1^2 + N_2^2$ pulses of the state ρ_2 with the phase basis, and N_2^3 pulses of the state ρ_3 with the phase basis, respectively. Similarly, Eve is assumed to control the rates \bar{b}_\times^1 , \bar{b}_\times^2 , and \bar{b}_\times^3 that Bob detects and phase error occurs among $N_1^1 + N_2^1 + N^1 + N^{1'}$ pulses, $N_1^2 + N_2^2$ pulses, and N_2^3 pulses, respectively. In the following discussion, we use the parameters $\bar{a}_\times^1 := \bar{q}^1 - \bar{b}_\times^1$, $\bar{a}_\times^2 := \bar{q}_\times^2 - \bar{b}_\times^2$, $\bar{a}_\times^3 := \bar{q}_\times^3 - \bar{b}_\times^3$, instead of \bar{q}^1 , \bar{q}_\times^2 , \bar{q}_\times^3 . For a simplicity, we employ the notations $\bar{\mathbf{a}} := (\bar{a}_\times^1, \bar{a}_\times^2, \bar{a}_\times^3)$ and $\bar{\mathbf{b}} := (\bar{b}_\times^1, \bar{b}_\times^2, \bar{b}_\times^3)$. Eve is also assumed to control the parameters \bar{q}^0 , $\bar{\mathbf{a}}$ and $\bar{\mathbf{b}}$ depending on the breakdown of the total $N_0 + N_1 + N_2 + N + N'$ pulses. Further, Eve is assumed to choose these values stochastically. Hence, the simultaneous distribution with condition for $\vec{\mathbf{N}}$ can be written as $Q_e(\bar{q}^0, \bar{\mathbf{a}}, \bar{\mathbf{b}} | \vec{\mathbf{N}})$. Hence, we have to analyze the security for respective breakdown of the total $N_0 + N_1 + N_2 + N + N'$ pulses.

5. Preparation concerning behavior of random variables

First, we summarize fundamental knowledge concerning behavior of random variables. When the true distribution is the N -trial binary distribution with success probability p , which is denoted by $\text{Bin}(N, p)$, we also denote the upper percent point with probability α by $Y^+(N, p, \alpha)$, and denote the lower percent point with probability α by $Y^-(N, p, \alpha)$. Then, we define $p^+(N, p, \alpha) := Y^+(N, p, \alpha)/N$, and $p^-(N, p, \alpha) := Y^-(N, p, \alpha)/N$. When we observe the value k subject to the binomial distribution $\text{Bin}(N, p)$ with the trial N and probability p , we denote the lower confidence limit of the lower one-sided interval estimation with the confidential level $1 - \alpha$ by $p^-(N, k, \alpha)$. Similarly, we denote the upper confidence limit of the upper one-sided interval estimation with the confidential level $1 - \alpha$ by $p^+(N, k, \alpha)$. Then, we define $X^-(N, k, \alpha) := p^-(N, k, \alpha)N$, and $X^+(N, k, \alpha) := p^+(N, k, \alpha)N$.

Next, we study the stochastic behavior of the measured values $\mathbf{M} = (M, M_0, M_1, M_2, M_3)$ under the assumption that the parameters $\bar{q}^0, \bar{\mathbf{a}}, \bar{\mathbf{b}}$ are unknown, but is fixed to certain values. The number of vacuum pulses is $N_0 + N_1^0 + N_2^0$. The detection ratio in Bob's side among $N_0 + N_1^0 + N_2^0$ vacuum pulses is fixed. N_0 vacuum pulses are randomly chosen from $N_0 + N_1^0 + N_2^0$ vacuum pulses, and M_0 is the number of detected pulses among these N_0 vacuum pulses. Hence, the random variable M_0 obeys the hypergeometric distribution. When $R > \bar{q}^0$, the probability $\Pr\{\frac{M_0}{N_0} > R\}$ under the above hypergeometric distribution is smaller than that under the binomial distribution $\text{Bin}(N_0, \bar{q}^0)$. This is because, if the detection rate among the initial L pulses is greater than R , the detecting probability of the $L + 1$ -th pulse is less than \bar{q}^0 for $L \leq N_0$ in the case of hypergeometric distribution.

Thus, we obtain

$$\Pr_{\mathbf{M}, \mathbf{J} | \bar{q}^0, \bar{\mathbf{a}}, \bar{\mathbf{b}}, \bar{\mathbf{N}}} \{M_0 < Y^-(N_0, \bar{q}^0, \epsilon)\} \leq \epsilon \quad (15)$$

$$\Pr_{\mathbf{M}, \mathbf{J} | \bar{q}^0, \bar{\mathbf{a}}, \bar{\mathbf{b}}, \bar{\mathbf{N}}} \{M_0 > Y^+(N_0, \bar{q}^0, \epsilon)\} \leq \epsilon, \quad (16)$$

where $\Pr_{\mathbf{M}, \mathbf{J} | \bar{q}^0, \bar{\mathbf{a}}, \bar{\mathbf{b}}, \bar{\mathbf{N}}}$ is the distribution concerning the random variables \mathbf{M}, \mathbf{J} when $\bar{q}^0, \bar{\mathbf{a}}, \bar{\mathbf{b}}, \bar{\mathbf{N}}$ are fixed. That is, we obtain

$$\Pr_{\mathbf{M}, \mathbf{J} | \bar{q}^0, \bar{\mathbf{a}}, \bar{\mathbf{b}}, \bar{\mathbf{N}}} \{\bar{M}_0 < X^-(N_0, M_0, \epsilon)\} \leq \epsilon \quad (17)$$

$$\Pr_{\mathbf{M}, \mathbf{J} | \bar{q}^0, \bar{\mathbf{a}}, \bar{\mathbf{b}}, \bar{\mathbf{N}}} \{\bar{M}_0 > X^+(N_0, M_0, \epsilon)\} \leq \epsilon, \quad (18)$$

where \bar{M}_0 is the expectation of M_0 , which equals $\bar{q}^0 N_0$.

Next, we focus on N_1 pulses with intensity μ_1 , which contain N_1^0 vacuum pulses, N_1^1 pulses with the single-photon state, and N_1^2 pulses with the state ρ_2 . Assume that $N_1^2 = 0$, $\frac{\bar{q}^0}{2} < \bar{a}_\times^1$, and \bar{M}_1 is the expectation of M_1 , which equals $\frac{\bar{q}^0}{2} N_1^0 + \bar{a}_\times^1 N_1^1 + \bar{a}_\times^2 N_1^2$. The probability $\Pr\{\frac{M_1}{N_1} > R\}$ for $R > \bar{M}_1/N_1$ is smaller than that when M_1 obeys the binomial distribution $\text{Bin}(N_1, \bar{M}_1/N_1)$. This fact can be shown as follows. Assume that the detection rate among the initial L pulses is greater than R . Under the above condition, the ratio of the single-photon pulses among initial L pulses is higher than $\frac{N_1^1}{N_1}$ with probability more than $1/2$. Conversely, under the above condition, the ratio

of the single-photon pulses among remaining $N_1 - L$ pulses is smaller than $\frac{N_1^1}{N_1}$ with probability more than $1/2$. Hence, the detecting probability of the $L + 1$ -th pulse is less than \bar{M}_1/N_1 . Therefore,

$$\Pr_{\mathbf{M}, \mathbf{J} | \bar{q}^0, \bar{a}, \bar{b}, \bar{\mathbf{N}}} \{M_1 < Y^-(N_1, \frac{\bar{M}_1}{N_1}, \epsilon)\} \leq \epsilon \quad (19)$$

$$\Pr_{\mathbf{M}, \mathbf{J} | \bar{q}^0, \bar{a}, \bar{b}, \bar{\mathbf{N}}} \{M_1 > Y^+(N_1, \frac{\bar{M}_1}{N_1}, \epsilon)\} \leq \epsilon, \quad (20)$$

which implies that

$$\Pr_{\mathbf{M}, \mathbf{J} | \bar{q}^0, \bar{a}, \bar{b}, \bar{\mathbf{N}}} \{\bar{M}_1 < X^-(N_1, M_1, \epsilon)\} \leq \epsilon \quad (21)$$

$$\Pr_{\mathbf{M}, \mathbf{J} | \bar{q}^0, \bar{a}, \bar{b}, \bar{\mathbf{N}}} \{\bar{M}_0 > X^+(N_1, M_1, \epsilon)\} \leq \epsilon. \quad (22)$$

Repeating a similar discussion, we can show the above relations without the condition $N_1^2 = 0$.

Similarly, the expectations of M_2 and M_3 are calculated to $\frac{\bar{q}^0}{2}N_2^0 + \bar{a}_\times^1 N_2^1 + \bar{a}_\times^2 N_2^2 + \bar{a}_\times^3 N_2^3$ and $\frac{\bar{q}^0}{2}N_1^0 + \bar{b}_\times^1 N_1^1 + \bar{b}_\times^2 N_1^2$, and are denoted by \bar{M}_2 and \bar{M}_3 , respectively. Then, we obtain

$$\Pr_{\mathbf{M}, \mathbf{J} | \bar{q}^0, \bar{a}, \bar{b}, \bar{\mathbf{N}}} \{\bar{M}_2 < X^-(N_2, M_2, \epsilon)\} \leq \epsilon \quad (23)$$

$$\Pr_{\mathbf{M}, \mathbf{J} | \bar{q}^0, \bar{a}, \bar{b}, \bar{\mathbf{N}}} \{\bar{M}_2 > X^+(N_2, M_2, \epsilon)\} \leq \epsilon \quad (24)$$

$$\Pr_{\mathbf{M}, \mathbf{J} | \bar{q}^0, \bar{a}, \bar{b}, \bar{\mathbf{N}}} \{\bar{M}_3 < X^-(N_1, M_3, \epsilon)\} \leq \epsilon \quad (25)$$

$$\Pr_{\mathbf{M}, \mathbf{J} | \bar{q}^0, \bar{a}, \bar{b}, \bar{\mathbf{N}}} \{\bar{M}_3 > X^+(N_1, M_3, \epsilon)\} \leq \epsilon. \quad (26)$$

For a detail discussion, see [40].

6. Outlines of derivation of sacrifice bit length and security proof

In this section, we give the outlines of derivation of sacrifice bit length and its security proof while their details will be given in latter sections. In the following, we treat the case when the intensity of the signal pulse (the signal intensity) is μ_1 and the intensity of the decoy pulse (the decoy intensity) is μ_2 .

First, we fix the breakdown $\bar{\mathbf{N}}$ of transmitted pulses. Then, we will give the sacrifice bit length in the privacy amplification as a function of the measured values $\mathbf{M} = (M, M_0, M_1, M_2, M_3)$ and the breakdown $\bar{\mathbf{N}}$. For this purpose, we introduce two conditions for the breakdown $\bar{\mathbf{N}}$.

Condition 1

$$N_1^1 N_2^2 - N_2^1 N_1^2 > 0. \quad (27)$$

Condition 2

$$N_1^2 N_2^0 - N_1^0 N_2^2 < 0. \quad (28)$$

Condition 3

$$-\frac{N_1^2 N_2^0 - N_2^2 N_1^0}{2(N_1^1 N_2^2 - N_1^2 N_2^1)} - \frac{N_1^0}{N_1^1} < 0, \quad (29)$$

When the all values take their expectation, the left hand side of Condition 1 is $\frac{1}{2}e^{-\mu_1-\mu_2}\mu_1\mu_2(\mu_2-\mu_1)\omega_2N_1N_2$, and is positive. In the same assumption, the left hand side of Condition 2 is $\frac{1}{2}N_1N_2e^{-(\mu_1+\mu_2)}\omega_2(\mu_1^2-\mu_2^2)$, and is negative.

Condition 3 is equivalent with

$$N_2^2 - \frac{N_1^2 N_2^0}{N_1^0} = \frac{N_2^2 N_1^0 - N_1^2 N_2^0}{N_1^0} < \frac{2(N_1^1 N_2^2 - N_1^2 N_2^1)}{N_1^1} = 2N_2^2 - 2\frac{N_1^2 N_2^1}{N_1^1}. \quad (30)$$

The above condition is equivalent with

$$2\frac{N_1^2 N_2^1}{N_1^1} - \frac{N_1^2 N_2^0}{N_1^0} < N_2^2. \quad (31)$$

Then, this condition is converted to

$$2\frac{N_2^1}{N_1^1} < \frac{N_2^0}{N_1^0} + \frac{N_2^2}{N_1^2}. \quad (32)$$

When the all values take their expectation, the left hand side is $2\frac{\mu_2}{\mu_1}e^{-\mu_2+\mu_1}$, and the right hand side is $(\frac{\mu_2^2}{\mu_1^2} + 1)e^{-\mu_2+\mu_1}$. Then, the above condition holds. Hence, these three assumptions are natural.

We also assume a condition for the estimate \hat{a}_\times^1 and \hat{b}_\times^1 for \bar{a}_\times^1 and \bar{b}_\times^1 as follows.

Condition 4

$$\frac{\hat{b}_\times^1}{\hat{a}_\times^1 + \hat{b}_\times^1} \leq \frac{1}{8}. \quad (33)$$

Hence, in Section 7, for a given real number $\beta > 0$, we will give an upper bound of leaked information $\hat{\phi}_2(\mathbf{M}, \vec{\mathbf{N}})$ as a function of $\vec{\mathbf{N}}$ and measured values $\mathbf{M} = (M, M_0, M_1, M_2, M_3)$ satisfying the following.

$$\Pr_{\mathbf{M}, \mathbf{J} | \vec{q}^0, \vec{a}, \vec{b}, \vec{\mathbf{N}}} \{ \hat{\phi}_2(\mathbf{M}, \vec{\mathbf{N}}) < \phi(\mathbf{J}) \} \leq 14 \cdot 2^{-2\beta-8}. \quad (34)$$

This equation implies the relations

$$\begin{aligned} & \Pr_{\mathbf{M}, \mathbf{J} | \vec{\mathbf{N}}} \{ \hat{\phi}_2(\mathbf{M}, \vec{\mathbf{N}}) < \phi(\mathbf{J}) \} \\ &= \sum_{\vec{q}^0, \vec{a}, \vec{b}} Q_e(\vec{q}^0, \vec{a}, \vec{b} | \vec{\mathbf{N}}) \Pr_{\mathbf{M}, \mathbf{J} | \vec{q}^0, \vec{a}, \vec{b}, \vec{\mathbf{N}}} \{ \hat{\phi}_2(\mathbf{M}, \vec{\mathbf{N}}) < \phi(\mathbf{J}) \} \leq 14 \cdot 2^{-2\beta-8}, \end{aligned}$$

where Q_e is the conditional distribution of $\vec{q}^0, \vec{a}, \vec{b}$ with condition for the breakdown $\vec{\mathbf{N}}$.

Next, we define the set Ω_1 as the set of $\vec{\mathbf{N}}$ satisfying

$$N^0 \in [Y_1^-(N, e^{-\mu_1}, 2^{-2\beta-8}), Y_1^+(N, e^{-\mu_1}, 2^{-2\beta-8})] \quad (35)$$

$$N^1 \in [Y_1^-(N, \mu_1 e^{-\mu_1}, 2^{-2\beta-8}), Y_1^+(N, \mu_1 e^{-\mu_1}, 2^{-2\beta-8})] \quad (36)$$

$$N_1^0 \in [Y_1^-(N_1, e^{-\mu_1}, 2^{-2\beta-8}), Y_1^+(N_1, e^{-\mu_1}, 2^{-2\beta-8})] \quad (37)$$

$$N_1^1 \in [Y_1^-(N_1, \mu_1 e^{-\mu_1}, 2^{-2\beta-8}), Y_1^+(N_1, \mu_1 e^{-\mu_1}, 2^{-2\beta-8})] \quad (38)$$

$$N_2^0 \in [Y_1^-(N_2, e^{-\mu_2}, 2^{-2\beta-8}), Y_1^+(N_2, e^{-\mu_2}, 2^{-2\beta-8})] \quad (39)$$

$$N_2^1 \in [Y_1^-(N_2, \mu_2 e^{-\mu_2}, 2^{-2\beta-8}), Y_1^+(N_2, \mu_2 e^{-\mu_2}, 2^{-2\beta-8})] \quad (40)$$

$$N_2^2 \in [Y_1^-(N_2, \omega_2 \mu_2^2 e^{-\mu_2}, 2^{-2\beta-8}), Y_1^+(N_2, \omega_2 \mu_2^2 e^{-\mu_2}, 2^{-2\beta-8})]. \quad (41)$$

Then, we introduce the following condition.

Condition 5

$$\begin{aligned}
& Y_1^-(N_1, \mu_1 e^{-\mu_1}, 2^{-2\beta-8}) Y_1^-(N_2, \omega_2 \mu_2^2 e^{-\mu_2}, 2^{-2\beta-8}) \\
& > (N_1 - Y_1^-(N_1, e^{-\mu_1}, 2^{-2\beta-8}) - Y_1^-(N_1, \mu_1 e^{-\mu_1}, 2^{-2\beta-8})) Y_1^+(N_2, \mu_2 e^{-\mu_2}, 2^{-2\beta-8}), \\
& \quad Y_1^-(N_2, \omega_2 \mu_2^2 e^{-\mu_2}, 2^{-2\beta-8}) Y_1^-(N_1, e^{-\mu_1}, 2^{-2\beta-8}) \\
& > (N_1 - Y_1^-(N_1, e^{-\mu_1}, 2^{-2\beta-8}) - Y_1^-(N_1, \mu_1 e^{-\mu_1}, 2^{-2\beta-8})) Y_1^+(N_2, e^{-\mu_2}, 2^{-2\beta-8}), \\
& \quad \frac{Y_1^-(N_2, \omega_2 \mu_2^2 e^{-\mu_2}, 2^{-2\beta-8})}{N_1 - Y_1^-(N_1, e^{-\mu_1}, 2^{-2\beta-8}) - Y_1^-(N_1, \mu_1 e^{-\mu_1}, 2^{-2\beta-8})} + \frac{Y_1^-(N_2, e^{-\mu_2}, 2^{-2\beta-8})}{Y_1^+(N_1, e^{-\mu_1}, 2^{-2\beta-8})} \\
& > \frac{2Y_1^+(N_2, \mu_2 e^{-\mu_2}, 2^{-2\beta-8})}{Y_1^-(N_1, \mu_1 e^{-\mu_1}, 2^{-2\beta-8})}.
\end{aligned}$$

This condition is equivalent with the condition that any element $\vec{N} \in \Omega_1$ satisfies Conditions 1, 2, and 3.

Under this condition, in Section 9, we will give the value $\hat{\phi}_4(\mathbf{M})$ as a function of the measured value \mathbf{M} , which does not depend on the breakdown \vec{N} . This value will be decided to satisfy the inequality

$$\{\hat{\phi}_4(\mathbf{M}) < \hat{\phi}_2(\mathbf{M}, \vec{N})\} \subset \Omega_1^c. \quad (42)$$

Then, we can show the following theorem.

Theorem 1 *When Condition 5 holds, we obtain*

$$\Pr_{\mathbf{J}, \mathbf{M}}\{\hat{\phi}_4(\mathbf{M}) < \phi(\mathbf{J})\} \leq 3 \cdot 2^{-2\beta-5}. \quad (43)$$

Proof. The above definition yields that

$$\Pr_{\vec{N}} \Omega_1^c \leq 14 \cdot 2^{-2\beta-8}.$$

Since

$$\begin{aligned}
& \{\hat{\phi}_4(\mathbf{M}) < \phi(\mathbf{J})\} \subset \{\hat{\phi}_2(\mathbf{M}, \vec{N}) < \phi(\mathbf{J})\} \cup \{\hat{\phi}_4(\mathbf{M}) < \hat{\phi}_2(\mathbf{M}, \vec{N})\} \\
& \subset \{\hat{\phi}_2(\mathbf{M}, \vec{N}) < \phi(\mathbf{J})\} \cup \Omega_1^c \subset (\{\hat{\phi}_2(\mathbf{M}, \vec{N}) < \phi(\mathbf{J})\} \cap \Omega_1) \cup \Omega_1^c,
\end{aligned}$$

we have

$$\begin{aligned}
& \Pr_{\mathbf{J}, \mathbf{M}}\{\hat{\phi}_2(\mathbf{M}) < \phi(\mathbf{J})\} \leq \Pr_{\mathbf{M}, \mathbf{J}, \vec{N}}(\{\hat{\phi}_2(\mathbf{M}, \vec{N}) < \phi(\mathbf{J})\} \cap \Omega_1) + \Pr_{\mathbf{M}, \mathbf{J}, \vec{N}} \Omega_1^c \\
& \leq 7 \cdot 2^{-2\beta-8} + 14 \cdot 2^{-2\beta-8} \leq 24 \cdot 2^{-2\beta-8} = 3 \cdot 2^{-2\beta-5},
\end{aligned}$$

which implies the desired argument. \square

Therefore, when $\rho_{A,E}$ is the final state with the sacrifice bit length

$$S(\mathbf{M}) := \hat{\phi}_4(\mathbf{M}) + 2\beta + 5, \quad (44)$$

(8) implies that

$$\begin{aligned}
& P_{ph} \\
& \leq 2^{-2\beta-5} \Pr\{\hat{\phi}_4(\mathbf{M}) + 2\beta + 5 < \phi(\mathbf{J}) + 2\beta + 5\}^c + \Pr\{\hat{\phi}_4(\mathbf{M}) + 2\beta + 5 < \phi(\mathbf{J}) + 2\beta + 5\} \\
& = 2^{-2\beta-5} + 3 \cdot 2^{-2\beta-5} = 2^{-2\beta-3}.
\end{aligned}$$

Thus, the relation (2) implies

$$\|\rho_{A,E} - \rho_{\text{ideal}}\|_1 \leq 2\sqrt{2}2^{(-2\beta-3)/2} = 2^{-\beta}. \quad (45)$$

When the signal intensity is μ_2 and the decoy intensity is μ_1 , the above arguments hold with modifying (142) and (143) in the following way.

$$N^{0'} \in [Y_1^-(N', e^{-\mu_2}, 2^{-2\beta-8}), Y_1^+(N', e^{-\mu_2}, 2^{-2\beta-8})] \quad (46)$$

$$N^{1'} \in [Y_1^-(N', \mu_2 e^{-\mu_2}, 2^{-2\beta-8}), Y_1^+(N', \mu_2 e^{-\mu_2}, 2^{-2\beta-8})]. \quad (47)$$

In summary, since Theorem 1 requires Condition 5, we need to choose the parameters μ_1 , μ_2 , N , N_0 , N_1 , and N_2 so that Condition 5 holds. That is, we need to choose sufficiently large integers N , N_0 , N_1 , and N_2 . Otherwise, we cannot apply Theorem 1, i.e., we cannot guarantee the security.

The latter sections give a derivation of the sacrifice bit length S as a function of β , μ_1 , μ_2 , N , N_0 , N_1 , N_2 , and \mathbf{M} . The derivation is summarized as Fig. 3. In order to apply interval estimation and percent point, we have to decide which upper or lower bound to be used in the respective steps.. These decisions will be treated based on derivatives for respective variables. Hence, the calculations of these derivatives are main issues in the latter sections.

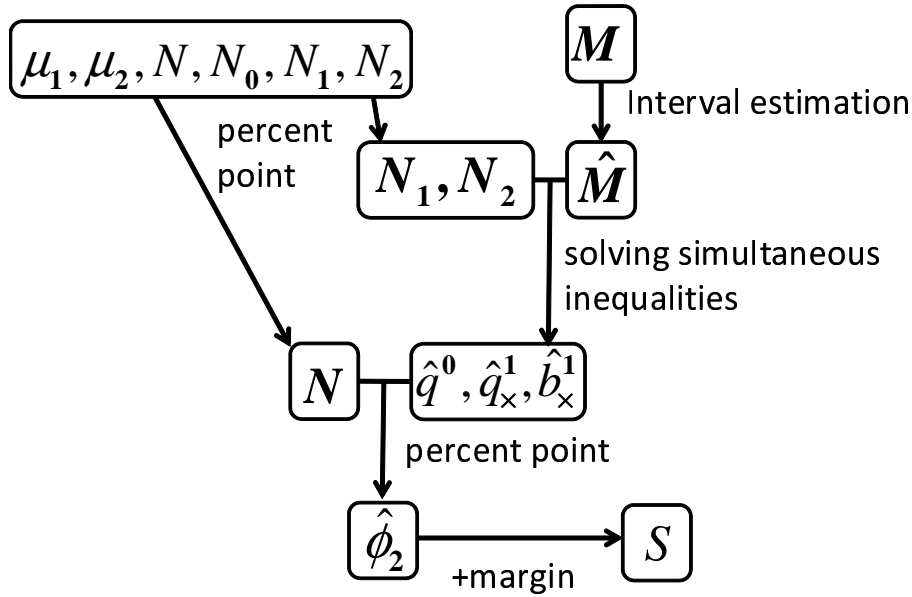


Figure 3. Outline of our derivation of the sacrifice bit length S . $\hat{\mathbf{M}} = (\hat{M}_0, \hat{M}_1, \hat{M}_2, \hat{M}_3)$ is the estimate of channel parameter $\bar{\mathbf{M}} = (\bar{M}_0, \bar{M}_1, \bar{M}_2, \bar{M}_3)$ given in Subsection 7.2.

7. Derivation of upper bound $\hat{\phi}_2$ of leaked information

7.1. Case when the channel parameters are given

First, we will derive an upper bound $\hat{\phi}_2$ of the leaked information ϕ when the channel parameters \bar{q}^0 , \bar{a}_\times^1 , \bar{b}_\times^1 and the breakdown \mathbf{N} of N pulses.

For this purpose, we describe the leaked information ϕ as a function of J^0 , J^1 , and $r^1 := J^3/J^1$:

$$\phi = M - J^0 - J^1(1 - h(\min\{r^1, 1/2\})). \quad (48)$$

That is, ϕ is monotone decreasing concerning J^0 and J^1 , and monotone increasing concerning r^1 . Hence, due to the same reason as (15), the channel parameters \bar{q}^0 , \bar{a}_\times^1 , and \bar{b}_\times^1 satisfy

$$\Pr_{J|\bar{q}^0, \bar{a}, \bar{b}, \mathbf{N}}\{J^0 \leq Y^-(N^0, \bar{q}^0, 2^{-2\beta-8})\} \leq 2^{-2\beta-8} \quad (49)$$

$$\Pr_{J|\bar{q}^0, \bar{a}, \bar{b}, \mathbf{N}}\{J^1 \leq Y^-(N^1, \bar{a}_\times^1 + \bar{b}_\times^1, 2^{-2\beta-8})\} \leq 2^{-2\beta-8}. \quad (50)$$

Using this fact, we give estimates of J^0 and J^1 as

$$\hat{J}^0(\bar{q}^0, N^0) := Y^-(N^0, \bar{q}^0, 2^{-2\beta-8}) \quad (51)$$

$$\hat{J}^1(\bar{a}_\times^1, \bar{b}_\times^1, N^1) := Y^-(N^1, \bar{a}_\times^1 + \bar{b}_\times^1, 2^{-2\beta-8}). \quad (52)$$

Since $p^+(j^1, \bar{a}_\times^1 + \bar{b}_\times^1, 2^{-2\beta-8}) \leq p^+(\tilde{j}^1, \bar{a}_\times^1 + \bar{b}_\times^1, 2^{-2\beta-8})$ holds for $j^1 \geq \tilde{j}^1$, (50) implies that

$$\begin{aligned} & \Pr_{J|\bar{q}^0, \bar{a}, \bar{b}, \mathbf{N}}\{p^+(J^1, \bar{a}_\times^1 + \bar{b}_\times^1, 2^{-2\beta-8}) \geq p^+(\hat{J}^1(\bar{a}_\times^1, \bar{b}_\times^1, N^1), \bar{a}_\times^1 + \bar{b}_\times^1, 2^{-2\beta-8})\} \\ &= \Pr_{J|\bar{q}^0, \bar{a}, \bar{b}, \mathbf{N}}\{J^1 \leq Y^-(N^1, \bar{a}_\times^1 + \bar{b}_\times^1, 2^{-2\beta-8})\} \leq 2^{-2\beta-8}. \end{aligned}$$

Using the relation

$$\Pr_{J|\bar{q}^0, \bar{a}, \bar{b}, \mathbf{N}}\{\frac{J^3}{J^1} \geq p^+(J^1, \bar{a}_\times^1 + \bar{b}_\times^1, 2^{-2\beta-8})\} \leq 2^{-2\beta-8}, \quad (53)$$

we obtain

$$\begin{aligned} & \{\frac{J^3}{J^1} \geq p^+(\hat{J}^1(\bar{a}_\times^1, \bar{b}_\times^1, N^1), \bar{a}_\times^1 + \bar{b}_\times^1, 2^{-2\beta-8})\} \\ & \subset \left(\{p^+(J^1, \bar{a}_\times^1 + \bar{b}_\times^1, 2^{-2\beta-8}) \geq p^+(\hat{J}^1(\bar{a}_\times^1, \bar{b}_\times^1, N^1), \bar{a}_\times^1 + \bar{b}_\times^1, 2^{-2\beta-8})\} \right. \\ & \quad \left. \cup \{\frac{J^3}{J^1} \geq p^+(J^1, \bar{a}_\times^1 + \bar{b}_\times^1, 2^{-2\beta-8})\} \right) \\ & \subset \{J^1 \leq Y^-(N^1, \bar{a}_\times^1 + \bar{b}_\times^1, 2^{-2\beta-8})\} \cup \{\frac{J^3}{J^1} \geq p^+(J^1, \bar{a}_\times^1 + \bar{b}_\times^1, 2^{-2\beta-8})\}. \end{aligned} \quad (54)$$

Therefore, we give an estimate of $r^1 := \frac{J^3}{J^1}$ by

$$\hat{r}_\times^1(\bar{a}_\times^1, \bar{b}_\times^1, N^1) := p^+(\hat{J}^1(\bar{a}_\times^1, \bar{b}_\times^1, N^1), \frac{\bar{b}_\times^1}{\bar{a}_\times^1 + \bar{b}_\times^1}, 2^{-2\beta-8}). \quad (55)$$

Using the above relations, we give an estimate of ϕ by

$$\hat{\phi}_1(\mathbf{N}, \bar{q}^0, \bar{a}_\times^1, \bar{b}_\times^1) := M - \hat{J}^0(\bar{q}^0, N^0) - \hat{J}^1(\bar{a}_\times^1, \bar{b}_\times^1, N^1)(1 - h(\min\{\hat{r}_\times^1(\bar{a}_\times^1, \bar{b}_\times^1, N^1), 1/2\})). \quad (56)$$

Due to (49), (50), (54), and (53), the estimate satisfies

$$\begin{aligned}
& \{\phi(\mathbf{J}) > \hat{\phi}_1(\mathbf{N}, \bar{q}^0, \bar{a}_\times^1, \bar{b}_\times^1)\} \\
& \subset \left(\{J^0 \leq Y^-(N^0, \bar{q}^0, 2^{-2\beta-8})\} \cup \{J^1 \leq Y^-(N^1, \bar{a}_\times^1, \bar{b}_\times^1, 2^{-2\beta-8})\} \right. \\
& \quad \left. \cup \left\{ \frac{J^3}{J^1} \geq p^+(\hat{J}^1(\bar{a}_\times^1, \bar{b}_\times^1, N^1), \frac{\bar{b}_\times^1}{\bar{a}_\times^1 + \bar{b}_\times^1}, 2^{-2\beta-8}) \right\} \right) \\
& \subset \left(\{J^0 \leq Y^-(N^0, \bar{q}^0, 2^{-2\beta-8})\} \cup \{J^1 \leq Y^-(N^1, \bar{a}_\times^1, \bar{b}_\times^1, 2^{-2\beta-8})\} \right. \\
& \quad \left. \cup \left\{ \frac{J^3}{J^1} \geq p^+(J^1, \frac{\bar{b}_\times^1}{\bar{a}_\times^1 + \bar{b}_\times^1}, 2^{-2\beta-8}) \right\} \right). \tag{57}
\end{aligned}$$

Hence, we obtain

$$\Pr\{\phi(\mathbf{J}) > \hat{\phi}_1(\mathbf{N}, \bar{q}^0, \bar{a}_\times^1, \bar{b}_\times^1)\} \leq 3 \cdot 2^{-2\beta-8}. \tag{58}$$

7.2. Estimation of channel parameters \bar{q}^0 , \bar{q}^1 , and \bar{r}_\times^1

Next, in order to treat an upper bound of leaked information ϕ , we will give estimates of channel parameters \bar{q}^0 , \bar{q}^1 , and \bar{r}_\times^1 based on the measured values \mathbf{M} and the breakdown $\vec{\mathbf{N}}$ of pulses. For this purpose, we treat the dependence of the leaked information $\hat{\phi}_1(\mathbf{N}, \bar{q}^0, \bar{a}_\times^1, \bar{b}_\times^1, 2^{-2\beta-8})$ concerning the channel parameters \bar{q}^0 , \bar{a}_\times^1 , and \bar{b}_\times^1 . That is, we evaluate the partial derivative of $\hat{\phi}_1(\mathbf{N}, \bar{q}^0, \bar{a}_\times^1, \bar{b}_\times^1)$ concerning \bar{q}^0 , \bar{a}_\times^1 , and \bar{b}_\times^1 .

However, while the partial derivatives of $\hat{J}^0(\bar{q}^0, N^0)$, $\hat{J}^1(\bar{a}_\times^1, \bar{b}_\times^1, N^1)$, and $\hat{r}_\times^1(\bar{a}_\times^1, \bar{b}_\times^1, N^1)$ are needed, their calculations are not easy. When N^0 and N^1 are sufficiently large, these values take the same values as $\bar{q}^0 N^0$, $(\bar{a}_\times^1 + \bar{b}_\times^1) N^1$, and $\frac{\bar{b}_\times^1}{\bar{a}_\times^1 + \bar{b}_\times^1}$, and the variations due to the fluctuations of \bar{q}^0 , \bar{a}_\times^1 , and \bar{b}_\times^1 are negligible.

Hence, we can regard the derivatives of $\hat{J}^0(\bar{q}^0, N^0)$, $\hat{J}^1(\bar{a}_\times^1, \bar{b}_\times^1, N^1)$, and $\hat{r}_\times^1(\bar{a}_\times^1, \bar{b}_\times^1, N^1)$ as the same as those of $\bar{q}^0 N^0$, $(\bar{a}_\times^1 + \bar{b}_\times^1) N^1$, and $\frac{\bar{b}_\times^1}{\bar{a}_\times^1 + \bar{b}_\times^1}$. Thus, we obtain

$$\frac{\partial \hat{J}^0(\bar{q}^0, N^0)}{\partial \bar{q}^0} = N^0 \tag{59}$$

$$\frac{\partial \hat{J}^1(\bar{a}_\times^1, \bar{b}_\times^1, N^1)}{\partial \bar{a}_\times^1} = N^1, \quad \frac{\partial \hat{J}^1(\bar{a}_\times^1, \bar{b}_\times^1, N^1)}{\partial \bar{b}_\times^1} = N^1 \tag{60}$$

$$\frac{\partial \hat{r}_\times^1(\bar{a}_\times^1, \bar{b}_\times^1, N^1)}{\partial \bar{a}_\times^1} = -\frac{\bar{b}_\times^1}{(\bar{a}_\times^1 + \bar{b}_\times^1)^2}, \quad \frac{\partial \hat{r}_\times^1(\bar{a}_\times^1, \bar{b}_\times^1, N^1)}{\partial \bar{b}_\times^1} = \frac{\bar{a}_\times^1}{(\bar{a}_\times^1 + \bar{b}_\times^1)^2}. \tag{61}$$

Under this assumption, we have

$$\frac{\partial \hat{\phi}_1(\mathbf{N}, \bar{q}^0, \bar{a}_\times^1, \bar{b}_\times^1)}{\partial \bar{q}^0} = \frac{\partial \phi}{\partial J^0} \frac{\partial \hat{J}^0}{\partial \bar{q}^0} = -N^0 < 0 \tag{62}$$

$$\frac{\partial \hat{\phi}_1(\mathbf{N}, \bar{q}^0, \bar{a}_\times^1, \bar{b}_\times^1)}{\partial \bar{a}_\times^1} = \frac{\partial \phi}{\partial J^1} \frac{\partial \hat{J}^1}{\partial \bar{a}_\times^1} + \frac{\partial \phi}{\partial \bar{r}_\times^1} \frac{\partial \hat{r}_\times^1}{\partial \bar{a}_\times^1} = -N^1 \left(1 + \log \frac{\bar{a}_\times^1}{\bar{a}_\times^1 + \bar{b}_\times^1}\right) < 0 \tag{63}$$

$$\frac{\partial \hat{\phi}_1(\mathbf{N}, \bar{q}^0, \bar{a}_\times^1, \bar{b}_\times^1)}{\partial \bar{b}_\times^1} = \frac{\partial \phi}{\partial J^1} \frac{\partial \hat{J}^1}{\partial \bar{b}_\times^1} + \frac{\partial \phi}{\partial \bar{r}_\times^1} \frac{\partial \hat{r}_\times^1}{\partial \bar{b}_\times^1} = -N^1 \left(1 + \log \frac{\bar{b}_\times^1}{\bar{a}_\times^1 + \bar{b}_\times^1}\right) > 0 \tag{64}$$

because $\frac{\bar{b}_\times^1}{\bar{a}_\times^1 + \bar{b}_\times^1} < \frac{1}{2}$. Therefore, smaller \bar{q}^0 and \bar{a}_\times^1 yield larger $\hat{\phi}_1$, and larger \bar{b}_\times^1 yields larger $\hat{\phi}_1$. That is, it is needed to estimate \bar{q}^0 and \bar{a}_\times^1 to be smaller and \bar{b}_\times^1 to be larger.

In the following, we treat the estimation of \bar{q}^0 , \bar{a}_\times^1 , and \bar{b}_\times^1 . Here, we should remark that we have two kinds of channel parameters. The first kind of parameters are \bar{q}^0 , \bar{a}_\times^1 , and \bar{b}_\times^1 , which are directly linked to the eavesdropping and cannot be measured directly. The other kind of parameters are the detection rate $p_{1,\times}$ and $p_{2,\times}$ of pulses of the phase basis with intensities μ_1 and μ_2 , which can be measured directly. Similarly, as other latter kind of parameters, we have the rates $s_{1,\times}$ and $s_{2,\times}$ that Bob detects and the error occurs of the phase basis with intensities μ_1 and μ_2 , which also can be measured directly.

The expectations \bar{M}_0 , \bar{M}_1 , \bar{M}_2 , and \bar{M}_3 of M_0 , M_1 , M_2 , and M_3 have characterizations $\bar{M}_0 = p_0 N_0$, $\bar{M}_1 = (p_{1,\times} - s_{1,\times}) N_1$, $\bar{M}_2 = (p_{2,\times} - s_{2,\times}) N_2$, and $\bar{M}_3 = s_{1,\times} N_1$. Hence, we can regard \bar{M}_0 , \bar{M}_1 , \bar{M}_2 , and \bar{M}_3 as the second kind of channel parameters. In our setting, \bar{M}_0 , \bar{M}_1 , \bar{M}_2 , and \bar{M}_3 are easier to treat than p_0 , $p_{1,\times}$, $p_{2,\times}$, $s_{1,\times}$, and $s_{2,\times}$. Thus, in the following, we estimate an upper bound of the leaked information ϕ via the estimation of the channel parameters \bar{q}^0 , \bar{a}_\times^1 , and \bar{b}_\times^1 when the channel parameters $\bar{\mathbf{M}} := (\bar{M}_0, \bar{M}_1, \bar{M}_2, \bar{M}_3)$ and the breakdown $\vec{\mathbf{N}}$ of pulses are given. Then, using the expansion formula (13), we obtain

$$\bar{M}_0 = \bar{q}^0 N_0 \quad (65)$$

$$\bar{M}_1 = \frac{\bar{q}^0}{2} N_1^0 + \bar{a}_\times^1 N_1^1 + \bar{a}_\times^2 N_1^2 \quad (66)$$

$$\bar{M}_2 = \frac{\bar{q}^0}{2} N_2^0 + \bar{a}_\times^1 N_2^1 + \bar{a}_\times^2 N_2^2 + \bar{a}_\times^3 N_2^3, \quad (67)$$

which implies the matrix equation

$$\begin{pmatrix} \bar{M}_0 \\ \bar{M}_1 \\ \bar{M}_2 - \bar{a}_\times^3 N_2^3 \end{pmatrix} = \begin{pmatrix} N_0 & 0 & 0 \\ N_1^0/2 & N_1^1 & N_1^2 \\ N_2^0/2 & N_2^1 & N_2^2 \end{pmatrix} \begin{pmatrix} \bar{q}^0 \\ \bar{a}_\times^1 \\ \bar{a}_\times^2 \end{pmatrix}. \quad (68)$$

Solving the above, we obtain

$$\bar{q}^0 = \hat{q}^0(\bar{\mathbf{M}}) := \frac{\bar{M}_0}{N_0} \quad (69)$$

$$\bar{a}_\times^1 = \hat{a}_\times^1(\bar{\mathbf{M}}, \vec{\mathbf{N}}) + A_1 \bar{a}_\times^3 \quad (70)$$

$$\hat{a}_\times^1(\bar{\mathbf{M}}, \vec{\mathbf{N}}) := \frac{N_2^2(\bar{M}_1 - \bar{M}_0 N_1^0/2N_0) - N_1^2(\bar{M}_2 - \bar{M}_0 N_2^0/2N_0)}{N_1^1 N_2^2 - N_2^1 N_1^2} \quad (71)$$

$$A_1 := \frac{N_1^2 N_2^3}{N_1^1 N_2^2 - N_2^1 N_1^2}. \quad (72)$$

Since A_1 and \bar{a}_\times^3 are non-negative, we obtain a lower bound of \bar{a}_\times^1 .

$$\bar{a}_\times^1 \geq \hat{a}_\times^1(\bar{\mathbf{M}}, \vec{\mathbf{N}}). \quad (73)$$

Similarly, we have

$$\bar{M}_3 = \bar{q}^0 N_1^0/2 + \bar{b}_\times^1 N_1^1 + \bar{b}_\times^2 N_1^2 \quad (74)$$

Since \bar{b}_\times^2 is non-negative, we obtain an upper bound of \bar{b}_\times^1 as

$$\bar{b}_\times^1 \leq \hat{b}_\times^1(\bar{\mathbf{M}}, \bar{\mathbf{N}}) := \frac{\bar{M}_3 - \frac{\bar{M}_0}{2N_0}N_1^0}{N_1^1}. \quad (75)$$

Then, we can define an upper bound of $\hat{\phi}_1$ as

$$\hat{\phi}_3(\hat{\mathbf{M}}(\mathbf{M}), \bar{\mathbf{N}}) := \hat{\phi}_1(\mathbf{N}, \hat{q}^0(\bar{M}_0), \hat{a}_\times^1(\hat{\mathbf{M}}(\mathbf{M}), \bar{\mathbf{N}}), \hat{b}_\times^1(\hat{\mathbf{M}}(\mathbf{M}), \bar{\mathbf{N}})), \quad (76)$$

where $\hat{\mathbf{M}}(\mathbf{M})$ is the estimate of $\bar{\mathbf{M}}$ when \mathbf{M} is observed.

Indeed, when $\bar{a}_\times^1 = \hat{a}_\times^1(\bar{\mathbf{M}}, \bar{\mathbf{N}})$ and $\bar{b}_\times^1 = \hat{b}_\times^1(\bar{\mathbf{M}}, \bar{\mathbf{N}})$, the relations

$$\bar{M}_1 = \frac{\bar{M}_0}{2N_0}N_1^0 + \hat{a}_\times^1 N_1^1 + \hat{a}_\times^2 N_1^2 \quad (77)$$

$$\bar{M}_2 = \frac{\bar{M}_0}{2N_0}N_2^0 + \hat{a}_\times^1 N_2^1 + \hat{a}_\times^2 N_2^2 \quad (78)$$

$$\bar{M}_3 = \frac{\bar{M}_0}{2N_0}N_1^0 + \hat{b}_\times^1 N_1^1 \quad (79)$$

hold.

Remark 1 When we extend the existing method[10, 8, 11] to our finite length setting, we obtain the following evaluation. In this case, we employ the parameter \bar{q}^1 instead of \bar{a}_\times^1 . Because smaller \bar{q}^1 yields larger $\hat{\phi}_1$, similar to \bar{a}_\times^1 , \bar{q}^1 can be estimated as

$$\hat{q}^1(\bar{\mathbf{M}}, \bar{\mathbf{N}}) := \frac{N_2^2(\bar{M}_1 + \bar{M}_3 - \bar{M}_0 N_1^0/N_0) - N_1^2(\bar{M}_2 + \bar{M}_4 - \bar{M}_0 N_2^0/N_0)}{N_1^1 N_2^2 - N_2^1 N_1^2}. \quad (80)$$

Then, in the upper bound $\hat{\phi}_1$ of the sacrifice bit size, $\hat{a}_\times^1 + \hat{b}_\times^1$ is replaced by $\hat{q}^1(\bar{\mathbf{M}}, \bar{\mathbf{N}})$. That is, we obtain an upper bound

$$\hat{\phi}_1(\mathbf{N}, \hat{q}^0(\bar{M}_0), \hat{q}^1(\bar{\mathbf{M}}, \bar{\mathbf{N}}) - \hat{b}_\times^1(\bar{\mathbf{M}}, \bar{\mathbf{N}}), \hat{b}_\times^1(\bar{\mathbf{M}}, \bar{\mathbf{N}})), \quad (81)$$

which is larger than $\hat{\phi}_3(\hat{\mathbf{M}}(\mathbf{M}), \bar{\mathbf{N}})$ because $\hat{q}^1(\bar{\mathbf{M}}, \bar{\mathbf{N}}) - \hat{b}_\times^1(\bar{\mathbf{M}}, \bar{\mathbf{N}}) \leq \hat{a}_\times^1(\bar{\mathbf{M}}, \bar{\mathbf{N}})$.

7.3. Estimation of another kind of channel parameters $\bar{\mathbf{M}}$

In this subsection, we treat the estimation of the channel parameters $\bar{\mathbf{M}}$ that is required to estimate the channel parameters \bar{q}^0 , \bar{a}_\times^1 , and \bar{b}_\times^1 when the breakdown $\bar{\mathbf{N}} = (\mathbf{N}, N_1, N_2)$ of pulses are known. That is, we consider the method to estimate \bar{M}_0 , \bar{M}_1 , \bar{M}_2 , and \bar{M}_3 from the measured value M_0, M_1, M_2 , and M_3 .

The partial derivatives of $\hat{q}^0(\bar{M}_0)$, $\hat{a}_\times^1(\bar{\mathbf{M}}, \bar{\mathbf{N}})$, and $\hat{b}_\times^1(\bar{\mathbf{M}}, \bar{\mathbf{N}})$ are calculated to

$$\frac{\partial \hat{q}^0}{\partial \bar{M}_0} = \frac{1}{N_0} > 0 \quad (82)$$

$$\frac{\partial \hat{q}^0}{\partial \bar{M}_1} = \frac{\partial \hat{q}^0}{\partial \bar{M}_2} = \frac{\partial \hat{q}^0}{\partial \bar{M}_3} = \frac{\partial \hat{q}^0}{\partial \bar{M}_4} = 0, \quad (83)$$

$$\frac{\partial \hat{a}_\times^1}{\partial \bar{M}_0} = \frac{(N_1^2 N_2^0 - N_2^2 N_1^0)}{2N_0(N_1^1 N_2^2 - N_2^1 N_1^2)} \quad (84)$$

$$\frac{\partial \hat{a}_\times^1}{\partial \bar{M}_1} = \frac{N_2^2}{N_1^1 N_2^2 - N_2^1 N_1^2} \quad (85)$$

$$\frac{\partial \hat{a}_x^1}{\partial \bar{M}_2} = - \frac{N_1^2}{N_1^1 N_2^2 - N_2^1 N_1^2} \quad (86)$$

$$\frac{\partial \hat{a}_x^1}{\partial \bar{M}_3} = 0, \quad (87)$$

$$\frac{\partial \hat{b}_x^1}{\partial \bar{M}_0} = - \frac{N_1^0}{2N_0 N_1^1} < 0 \quad (88)$$

$$\frac{\partial \hat{b}_x^1}{\partial \bar{M}_3} = \frac{1}{N_1^1} > 0 \quad (89)$$

$$\frac{\partial \hat{b}_x^1}{\partial \bar{M}_1} = \frac{\partial \hat{b}_x^1}{\partial \bar{M}_2} = 0. \quad (90)$$

Hence, applying Conditions 1 and 2 to (84),(85),(86), we obtain

$$\frac{\partial \hat{a}_x^1}{\partial \bar{M}_0} < 0, \quad \frac{\partial \hat{a}_x^1}{\partial \bar{M}_1} > 0, \quad \frac{\partial \hat{a}_x^1}{\partial \bar{M}_2} < 0.$$

Thus, due to (62), (63), (64), and Conditions 3 and 4, we obtain

$$\begin{aligned} \frac{\partial \hat{\phi}_1}{\partial \bar{M}_1} &< 0, \quad \frac{\partial \hat{\phi}_1}{\partial \bar{M}_2} > 0, \quad \frac{\partial \hat{\phi}_1}{\partial \bar{M}_3} > 0, \\ \frac{\partial \hat{\phi}_1}{\partial \bar{M}_0} &= -\frac{N^0}{N_0} - \frac{N^1(N_1^2 N_2^0 - N_2^2 N_1^0)(1 + \log \frac{\hat{a}_x^1}{\hat{a}_x^1 + \hat{b}_x^1})}{2N_0(N_1^1 N_2^2 - N_1^2 N_2^1)} + \frac{N^1 N_1^0(1 + \log \frac{\hat{b}_x^1}{\hat{a}_x^1 + \hat{b}_x^1})}{2N_0 N_1^1} \\ &\leq -\frac{N^0}{N_0} - \frac{N^1(N_1^2 N_2^0 - N_2^2 N_1^0)}{2N_0(N_1^1 N_2^2 - N_1^2 N_2^1)} - \frac{N^1 N_1^0}{N_0 N_1^1} < 0. \end{aligned}$$

Hence, we need to estimate \bar{M}_0 and \bar{M}_1 to be smaller, and \bar{M}_2 and \bar{M}_3 to be larger.

In the following, we employ

$$\hat{M}_0(M_0) := X^-(N_0, M_0, 2^{-2\beta-8}), \quad (91)$$

$$\hat{M}_1(M_1) := X^-(N_1, M_1, 2^{-2\beta-8}), \quad (92)$$

$$\hat{M}_2(M_2) := X^+(N_2, M_2, 2^{-2\beta-8}), \quad (93)$$

$$\hat{M}_3(M_3) := X^+(N_1, M_3, 2^{-2\beta-8}), \quad (94)$$

as estimates of \bar{M}_0 , \bar{M}_1 , \bar{M}_2 , and \bar{M}_3 . That is, defining

$$\hat{\phi}_2(\mathbf{M}, \vec{\mathbf{N}}) := \hat{\phi}_3(\hat{\mathbf{M}}(\mathbf{M}), \vec{\mathbf{N}}),$$

we obtain the following theorem.

Theorem 2 *When the breakdown $\vec{\mathbf{N}}$ satisfies Conditions 1 and 2, the relation*

$$\Pr_{\mathbf{J}, \mathbf{M} | \vec{q}^0, \vec{a}, \vec{b}, \vec{\mathbf{N}}} \{ \hat{\phi}_2(\mathbf{M}, \vec{\mathbf{N}}) < \phi(\mathbf{J}) \} \leq 7 \cdot 2^{-2\beta-8}$$

holds.

Proof. The definition of $\hat{\phi}_3$ given in (76) yields

$$\begin{aligned} \{ \phi(\mathbf{J}) > \hat{\phi}_2(\mathbf{M}, \vec{\mathbf{N}}) \} &\subset \{ \phi(\mathbf{J}) > \hat{\phi}_3(\bar{\mathbf{M}}, \vec{\mathbf{N}}) \} \cup \{ \hat{\phi}_3(\bar{\mathbf{M}}, \vec{\mathbf{N}}) > \hat{\phi}_2(\mathbf{M}, \vec{\mathbf{N}}) \} \\ &\subset \{ \phi(\mathbf{J}) > \hat{\phi}_1(\mathbf{N}, \vec{q}^0, \vec{a}_x^1, \vec{b}_x^1) \} \cup \{ \hat{\phi}_3(\bar{\mathbf{M}}, \vec{\mathbf{N}}) > \hat{\phi}_2(\mathbf{M}, \vec{\mathbf{N}}) \}. \end{aligned} \quad (95)$$

Hence, the above calculation concerning the partial derivatives and Conditions 1 and 2 imply

$$\begin{aligned} & \{\hat{\phi}_3(\bar{\mathbf{M}}, \vec{\mathbf{N}}) > \hat{\phi}_2(\mathbf{M}, \vec{\mathbf{N}})\} \\ & \subset \{\bar{M}_0 < \hat{M}_0(M_0)\} \cup \{\bar{M}_1 < \hat{M}_1(M_1)\} \cup \{\bar{M}_2 > \hat{M}_2(M_2)\} \cup \{\bar{M}_3 > \hat{M}_3(M_3)\}. \end{aligned}$$

Thus, it follows from the relations (17), (21), (24), and (26), that

$$\Pr_{\mathbf{J}, \mathbf{M} | \bar{q}^0, \bar{\mathbf{a}}, \bar{\mathbf{b}}, \vec{\mathbf{N}}} \{\hat{\phi}_3(\bar{\mathbf{M}}, \vec{\mathbf{N}}) > \hat{\phi}_2(\mathbf{M}, \vec{\mathbf{N}})\} \leq 4 \cdot 2^{-2\beta-8}.$$

Using (95) and (58), we obtain

$$\Pr_{\mathbf{J}, \mathbf{M} | \bar{q}^0, \bar{\mathbf{a}}, \bar{\mathbf{b}}, \vec{\mathbf{N}}} \{\hat{\phi}_2(\mathbf{M}, \vec{\mathbf{N}}) < \phi(\mathbf{J})\} \leq 3 \cdot 2^{-2\beta-8} + 4 \cdot 2^{-2\beta-8} \leq 7 \cdot 2^{-2\beta-8}.$$

□

8. Asymptotic analysis

8.1. Asymptotic key generation formulas

In this section, we consider the asymptotic setting. Then, we assume that all of $\vec{\mathbf{N}}$ equals their expectations. The channel parameters p_0 , $p_{1,\times}$, $p_{2,\times}$, $s_{1,\times}$, and $s_{2,\times}$ satisfy

$$\begin{aligned} \hat{q}^0 &= p_0 \\ \hat{a}_\times^1 &= \min\{\hat{a}_\times^1(p_{1,\times} - s_{1,\times}, p_{2,\times} - s_{2,\times}, \mu_1, \mu_2), 0\} \end{aligned} \tag{96}$$

$$\hat{b}_\times^1 = \hat{b}_\times^1(s_{1,\times}, \mu_1), \tag{97}$$

where

$$\begin{aligned} \hat{a}_\times^1(p_1, p_2, \mu_1, \mu_2) &:= \frac{\omega_2(\mu_2^2 e^{-\mu_2}(p_1 - p_0 e^{-\mu_1}/2) - \mu_1^2 e^{-\mu_1}(p_2 - p_0 e^{-\mu_2}/2))}{\omega_2 e^{-\mu_1-\mu_2} \mu_1 \mu_2 (\mu_2 - \mu_1)} \\ &= \frac{\mu_2^2 e^{\mu_1}(p_1 - p_0 e^{-\mu_1}/2) - \mu_1^2 e^{\mu_2}(p_2 - p_0 e^{-\mu_2}/2)}{\mu_1 \mu_2 (\mu_2 - \mu_1)} \\ \hat{b}_\times^1(s_1, \mu_1) &:= \frac{s_1 e^{\mu_1} - p_0/2}{\mu_1}. \end{aligned}$$

When the signal intensity is μ_1 , the key generation rate per a raw key is

$$\frac{M - \phi - M\eta h(\frac{s_1}{p_1})}{M} = \frac{\mu_1 e^{-\mu_1}(\hat{a}_\times^1 + \hat{b}_\times^1)(1 - h(\frac{\hat{b}_\times^1}{\hat{a}_\times^1 + \hat{b}_\times^1})) + e^{-\mu_1}p_0 - p_{1,\times}\eta h(\frac{s_{1,\times}}{p_{1,\times}})}{p_{1,\times}}.$$

Similarly, when the signal intensity is μ_2 , the key generation rate per a raw key is

$$\frac{M - \phi - M\eta h(\frac{s_2}{p_2})}{M} = \frac{\mu_2 e^{-\mu_2}(\hat{a}_\times^1 + \hat{b}_\times^1)(1 - h(\frac{\hat{b}_\times^1}{\hat{a}_\times^1 + \hat{b}_\times^1})) + e^{-\mu_2}p_0 - p_{2,\times}\eta h(\frac{s_{2,\times}}{p_{2,\times}})}{p_{2,\times}}.$$

Hence, the key generation rates per pulse with the coincidence basis are

$$R_1 := \mu_1 e^{-\mu_1}(\hat{a}_\times^1 + \hat{b}_\times^1)(1 - h(\frac{\hat{b}_\times^1}{\hat{a}_\times^1 + \hat{b}_\times^1})) + e^{-\mu_1}p_0 - p_{1,\times}\eta h(\frac{s_{1,\times}}{p_{1,\times}}) \tag{98}$$

$$R_2 := \mu_2 e^{-\mu_2}(\hat{a}_\times^1 + \hat{b}_\times^1)(1 - h(\frac{\hat{b}_\times^1}{\hat{a}_\times^1 + \hat{b}_\times^1})) + e^{-\mu_2}p_0 - p_{2,\times}\eta h(\frac{s_{2,\times}}{p_{2,\times}}), \tag{99}$$

which is the same as the application of the GLLP formula [6]. On the other hand, the existing method[10, 9, 11] yields the key generation rates per pulse with the coincidence basis as

$$\tilde{R}_1 := \mu_1 e^{-\mu_1} \hat{q}^1 (1 - h(\frac{\hat{b}_{\times}^1}{\hat{q}^1})) + e^{-\mu_1} p_0 - p_{1,\times} \eta h(\frac{s_{1,\times}}{p_{1,\times}}) \quad (100)$$

$$\tilde{R}_2 := \mu_2 e^{-\mu_2} \hat{q}^1 (1 - h(\frac{\hat{b}_{\times}^1}{\hat{q}^1})) + e^{-\mu_2} p_0 - p_{2,\times} \eta h(\frac{s_{2,\times}}{p_{2,\times}}). \quad (101)$$

with

$$\hat{q}^1 := \frac{\omega_2(\mu_2^2 e^{-\mu_2}(p_{1,\times} - p_0 e^{-\mu_1}) - \mu_1^2 e^{-\mu_1}(p_{2,\times} - p_0 e^{-\mu_2}))}{\omega_2 e^{-\mu_1 - \mu_2} \mu_1 \mu_2 (\mu_2 - \mu_1)} \quad (102)$$

$$= \frac{\mu_2^2 e^{\mu_1}(p_{1,\times} - p_0 e^{-\mu_1}) - \mu_1^2 e^{\mu_2}(p_{2,\times} - p_0 e^{-\mu_2})}{\mu_1 \mu_2 (\mu_2 - \mu_1)}. \quad (103)$$

8.2. Case when true intensities coincide with our intents

In the following, we consider the case with no eavesdropper, i.e., the case when the true intensities coincide with our intent intensities. That is, we adopt the following model with parameters α and s :

$$p_{i,\times} = 1 - e^{-\alpha \mu_i} + p_0, \quad s_{i,\times} = s(1 - e^{-\alpha \mu_i}) + \frac{p_0}{2}, \quad (104)$$

which implies $p_{i,\times} - s_{i,\times} = (1 - s)(1 - e^{-\alpha \mu_i}) + \frac{p_0}{2}$. Hence, the relations

$$\begin{aligned} \hat{a}_{\times}^1 &= \hat{a}_{\times}^1(\mu_1, \mu_2) \\ &:= \frac{\mu_2^2 e^{\mu_1}((1 - s)(1 - e^{-\alpha \mu_1}) + \frac{p_0}{2}(1 - e^{-\mu_1})) - \mu_1^2 e^{\mu_2}((1 - s)(1 - e^{-\alpha \mu_2}) + \frac{p_0}{2}(1 - e^{-\mu_2}))}{\mu_1 \mu_2 (\mu_2 - \mu_1)} \\ &= (\mu_2 \mu_1) \frac{\frac{(1-s+p_0/2)e^{\mu_1} - (1-s)e^{(1-\alpha)\mu_1} - \frac{p_0}{2}}{\mu_1^2} - \frac{(1-s+p_0/2)e^{\mu_2} - (1-s)e^{(1-\alpha)\mu_2} - \frac{p_0}{2}}{\mu_2^2}}{\mu_2 - \mu_1} \end{aligned} \quad (105)$$

$$\hat{b}_{\times}^1 = \hat{b}_{\times}^1(\mu_1) := \frac{s(1 - e^{-\alpha \mu_1})e^{\mu_1} + \frac{p_0}{2}(e^{\mu_1} - 1)}{\mu_1} \quad (106)$$

$$\begin{aligned} \hat{q}^1 &= \hat{q}^1(\mu_1, \mu_2) \\ &:= \frac{\mu_2^2 e^{\mu_1}((1 - e^{-\alpha \mu_1}) + p_0(1 - e^{-\mu_1})) - \mu_1^2 e^{\mu_2}((1 - e^{-\alpha \mu_2}) + p_0(1 - e^{-\mu_2}))}{\mu_1 \mu_2 (\mu_2 - \mu_1)} \\ &= (\mu_2 \mu_1) \frac{\frac{(1+p_0)e^{\mu_1} - e^{(1-\alpha)\mu_1} - p_0}{\mu_1^2} - \frac{(1+p_0)e^{\mu_2} - e^{(1-\alpha)\mu_2} - p_0}{\mu_2^2}}{\mu_2 - \mu_1} \end{aligned} \quad (107)$$

hold. Substituting the above \hat{a}_{\times}^1 , \hat{b}_{\times}^1 into (98),(99), we obtain R_1 and R_2 as functions of μ_1 and μ_2 in the following way:

$$\begin{aligned} R_1(\mu_1, \mu_2) &:= \mu_1 e^{-\mu_1} (\hat{a}_{\times}^1(\mu_1, \mu_2) + \hat{b}_{\times}^1(\mu_1, \mu_2)) (1 - h(\frac{\hat{b}_{\times}^1(\mu_1, \mu_2)}{\hat{a}_{\times}^1(\mu_1, \mu_2) + \hat{b}_{\times}^1(\mu_1, \mu_2)})) \\ &\quad + e^{-\mu_1} p_0 - p_{1,\times} \eta h(\frac{s_{1,\times}}{p_{1,\times}}) \end{aligned} \quad (108)$$

$$\begin{aligned}
R_2(\mu_1, \mu_2) := & \mu_2 e^{-\mu_2} (\hat{a}_\times^1(\mu_1, \mu_2) + \hat{b}_\times^1(\mu_1, \mu_2)) (1 - h(\frac{\hat{b}_\times^1(\mu_1, \mu_2)}{\hat{a}_\times^1(\mu_1, \mu_2) + \hat{b}_\times^1(\mu_1, \mu_2)})) \\
& + e^{-\mu_2} p_0 - p_{2,\times} \eta h(\frac{s_{2,\times}}{p_{2,\times}}),
\end{aligned} \tag{109}$$

As is calculated in Figure 4, our key generation rate R_2 improves the existing key generation rate \tilde{R}_2 .

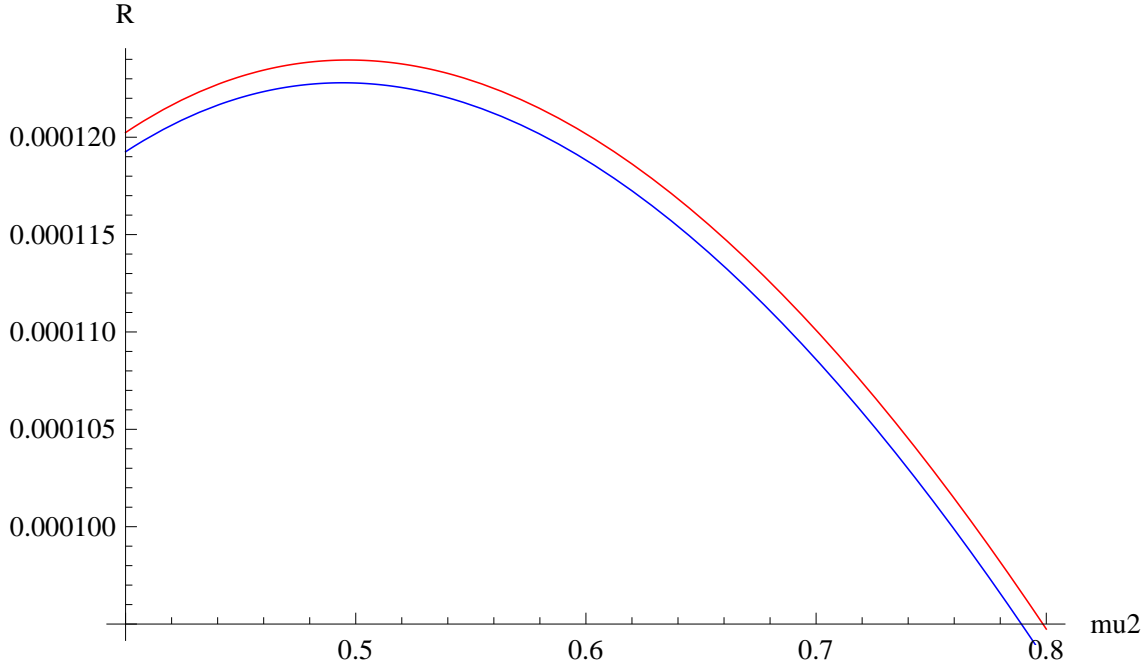


Figure 4. Our key generation rate R_2 (red) and existing key generation rate \tilde{R}_2 (blue) with $\tilde{\mu}_1 = 0.1$, $s = 0.03$, $p_0 = 4.0 \times 10^{-7}$, $\alpha = 1.0 \times 10^{-3}$.

Then, using (106) and the first equations of (105) and (107), we obtain

$$\lim_{\mu_1 \rightarrow +0} \hat{a}_\times^1(\mu_1, \mu_2) = (1-s)\alpha + \frac{p_0}{2}, \quad \lim_{\mu_1 \rightarrow +0} \hat{b}_\times^1(\mu_1) = s\alpha + \frac{p_0}{2}, \quad \lim_{\mu_1 \rightarrow +0} \hat{q}^1(\mu_1, \mu_2) = \alpha + p_0.$$

Further, the both relations in (105), and the derivative of $\frac{(1+p_0)e^{\mu_1} - e^{(1-\alpha)\mu_1} - p_0}{\mu_1^2}$ concerning μ_1 imply

$$\begin{aligned}
\hat{a}_\times^1(\mu_1) &:= \lim_{\mu_2 \rightarrow \mu_1 + 0} \hat{a}_\times^1(\mu_1, \mu_2) \\
&= \mu_1^2 \lim_{\mu_2 \rightarrow \mu_1 + 0} \frac{\frac{(1-s+p_0/2)e^{\mu_1} - (1-s)e^{(1-\alpha)\mu_1} - p_0/2}{\mu_1^2} - \frac{(1-s+p_0/2)e^{\mu_2} - (1-s)e^{(1-\alpha)\mu_2} - p_0/2}{\mu_2^2}}{\mu_2 - \mu_1} \\
&= 2 \frac{(1-s+p_0/2)e^{\mu_1} - (1-s)e^{(1-\alpha)\mu_1} - p_0/2}{\mu_1} - (1-s+p_0/2)e^{\mu_1} + (1-s)(1-\alpha)e^{(1-\alpha)\mu_1} \\
&= \frac{(2-\mu_1)(1-s+p_0/2) - (1-s)(2-(1-\alpha)\mu_1)e^{-\alpha\mu_1} - p_0e^{-\mu_1}}{\mu_1 e^{-\mu_1}} \\
\hat{q}^1(\mu_1) &:= \lim_{\mu_2 \rightarrow \mu_1 + 0} \hat{q}^1(\mu_1, \mu_2) = \mu_1^2 \lim_{\mu_2 \rightarrow \mu_1 + 0} \frac{\frac{(1+p_0)e^{\mu_1} - e^{(1-\alpha)\mu_1} - p_0}{\mu_1^2} - \frac{(1+p_0)e^{\mu_2} - e^{(1-\alpha)\mu_2} - p_0}{\mu_2^2}}{\mu_2 - \mu_1}
\end{aligned}$$

$$\begin{aligned}
&= 2 \frac{(1+p_0)e^{\mu_1} - e^{(1-\alpha)\mu_1} - p_0}{\mu_1} - (1+p_0)e^{\mu_1} + (1-\alpha)e^{(1-\alpha)\mu_1} \\
&= \frac{(2-\mu_1)(1+p_0) - (2-(1-\alpha)\mu_1)e^{-\alpha\mu_1} - 2p_0e^{-\mu_1}}{\mu_1 e^{-\mu_1}}.
\end{aligned}$$

Then, we obtain the following theorem.

Theorem 3 $(\hat{a}_\times^1(\mu_1, \mu_2) + \hat{b}_\times^1(\mu_1, \mu_2))(1 - h(\frac{\hat{b}_\times^1(\mu_1, \mu_2)}{\hat{a}_\times^1(\mu_1, \mu_2) + \hat{b}_\times^1(\mu_1, \mu_2)}))$ is monotone decreasing for μ_1 and μ_2 when $\frac{\hat{b}_\times^1(\mu_1, \mu_2)}{\hat{a}_\times^1(\mu_1, \mu_2) + \hat{b}_\times^1(\mu_1, \mu_2)} < \frac{1}{2}$.

Now, we fixed a signal intensity to be μ_s . Then, Theorem 3 implies

$$R_2(\mu_1, \mu_s) \geq R_2(\mu_2, \mu_s) \geq R_1(\mu_s, \mu_3) \geq R_1(\mu_s, \mu_4)$$

for $\mu_1 < \mu_2 < \mu_s < \mu_3 < \mu_4$. These inequalities imply that a smaller decoy intensity has a better key generation rate when the signal intensity is fixed.

Further, using Theorem 3, we obtain

$$\begin{aligned}
&\sup_{\mu_1: 0 < \mu_1 < \mu_2} (\hat{a}_\times^1(\mu_1, \mu_2) + \hat{b}_\times^1(\mu_1))(1 - h(\frac{\hat{b}_\times^1(\mu_1)}{\hat{a}_\times^1(\mu_1, \mu_2) + \hat{b}_\times^1(\mu_1)})) \\
&= \lim_{\mu_1 \rightarrow +0} (\hat{a}_\times^1(\mu_1, \mu_2) + \hat{b}_\times^1(\mu_1))(1 - h(\frac{\hat{b}_\times^1(\mu_1)}{\hat{a}_\times^1(\mu_1, \mu_2) + \hat{b}_\times^1(\mu_1)})) \\
&= (\alpha + p_0)(1 - h(\frac{\alpha s + \frac{p_0}{2}}{\alpha + p_0})), \tag{110}
\end{aligned}$$

$$\begin{aligned}
&\sup_{\mu_2: 0 < \mu_1 < \mu_2} (\hat{a}_\times^1(\mu_1, \mu_2) + \hat{b}_\times^1(\mu_1))(1 - h(\frac{\hat{b}_\times^1(\mu_1)}{\hat{a}_\times^1(\mu_1, \mu_2) + \hat{b}_\times^1(\mu_1)})) \\
&= \lim_{\mu_2 \rightarrow \mu_1 + 0} (\hat{a}_\times^1(\mu_1, \mu_2) + \hat{b}_\times^1(\mu_1))(1 - h(\frac{\hat{b}_\times^1(\mu_1)}{\hat{a}_\times^1(\mu_1, \mu_2) + \hat{b}_\times^1(\mu_1)})) \\
&= (\hat{a}_\times^1(\mu_1) + \hat{b}_\times^1(\mu_1))(1 - h(\frac{\hat{b}_\times^1(\mu_1)}{\hat{a}_\times^1(\mu_1) + \hat{b}_\times^1(\mu_1)})). \tag{111}
\end{aligned}$$

The right hand side of (110) equals the true parameter in the model (104). That is, when the signal pulse has the larger intensity μ_2 and we take the limit $\mu_1 \rightarrow 0$, our value (110) coincides with the value $(\hat{a}_\times^1 + \hat{b}_\times^1)h(\frac{\hat{b}_\times^1}{\hat{a}_\times^1 + \hat{b}_\times^1})$ with $\hat{a}_\times^1 = \frac{p_0}{2} + \alpha(1-s)$ and $\hat{b}_\times^1 = \frac{p_0}{2} + \alpha s$. Note that the true parameters \bar{a}_\times^1 and \bar{b}_\times^1 are $\frac{p_0}{2} + \alpha(1-s)$ and $\frac{p_0}{2} + \alpha s$.

These relations (110) and (111) imply the optimal key generation rate with a fixed signal intensity as (113) and (112). In particular, the case given in (110) yields the rate when we can perfectly estimate the parameters \bar{a}_\times^1 and \bar{b}_\times^1 as in (113). That is, the optimal rate coincides with the rate when the estimates \hat{a}_\times^1 and \hat{b}_\times^1 are the true parameters $\bar{a}_\times^1 = \frac{p_0}{2} + \alpha(1-s)$ and $\bar{b}_\times^1 = \frac{p_0}{2} + \alpha s$.

$$\begin{aligned}
R_1(\mu_1) &:= \sup_{\mu_2: 0 < \mu_1 < \mu_2} R_1(\mu_1, \mu_2) \\
&= \mu_1 e^{-\mu_1} (\hat{a}_\times^1(\mu_1) + \hat{a}_\times^1(\mu_1))(1 - h(\frac{\hat{b}_\times^1(\mu_1)}{\hat{a}_\times^1(\mu_1) + \hat{a}_\times^1(\mu_1)})) + e^{-\mu_1} p_0
\end{aligned}$$

$$- (1 - e^{-\alpha\mu_1} + p_0)\eta h\left(\frac{s(1 - e^{-\alpha\mu_1}) + \frac{p_0}{2}}{1 - e^{-\alpha\mu_1} + p_0}\right), \quad (112)$$

$$\begin{aligned} R_2(\mu_2) &:= \sup_{\mu_1: 0 < \mu_1 < \mu_2} R_2(\mu_1, \mu_2) \\ &= \mu_2 e^{-\mu_2}(\alpha + p_0)(1 - h(\frac{\alpha s + \frac{p_0}{2}}{\alpha + p_0})) + e^{-\mu_2}p_0 - (1 - e^{-\alpha\mu_2} + p_0)\eta h\left(\frac{s(1 - e^{-\alpha\mu_2}) + \frac{p_0}{2}}{1 - e^{-\alpha\mu_2} + p_0}\right) \end{aligned} \quad (113)$$

In particular, when $p_0 = 0$, we obtain

$$R_2(\mu_2) = \mu_2 e^{-\mu_2} \alpha (1 - h(s)) - (1 - e^{-\alpha\mu_2}) \eta h(s).$$

Proof. Define the function $f(a, b) := (a + b)(1 - h(\frac{b}{a+b}))$. Since

$$\frac{\partial f}{\partial a} = 1 + \log \frac{a}{a+b}, \quad \frac{\partial f}{\partial b} = 1 + \log \frac{b}{a+b},$$

it is sufficient to show that $\hat{a}_\times^1(\mu_1, \mu_2)$ is monotone decreasing for μ_1 and μ_2 and $\hat{b}_\times^1(\mu_1)$ is monotone increasing for μ_1 .

Since

$$\begin{aligned} \hat{b}_\times^1(\mu_1) &= \frac{s(e^{\mu_1} - e^{(1-\alpha)\mu_1}) + p_0(e^{\mu_1} - 1)/2}{\mu_1} \\ &= s\alpha + \frac{p_0}{2} + \sum_{n=2}^{\infty} (s(1 - (1-\alpha)^n) + \frac{p_0}{2}) \frac{\mu_1^{n-1}}{n!}, \end{aligned} \quad (114)$$

and $s(1 - (1-\alpha)^n) + \frac{p_0}{2} \geq 0$, $\hat{b}_\times^1(\mu_1)$ is monotone increasing for μ_1 .

Further,

$$\begin{aligned} \hat{a}_\times^1(\mu_1, \mu_2) &= (\mu_2 \mu_1) \frac{\frac{(1-s+p_0/2)e^{\mu_1} - (1-s)e^{(1-\alpha)\mu_1} - p_0/2}{\mu_1^2} - \frac{(1-s+p_0/2)e^{\mu_2} - (1-s)e^{(1-\alpha)\mu_2} - p_0/2}{\mu_2^2}}{\mu_2 - \mu_1} \\ &= (\mu_2 \mu_1) \frac{\sum_{n=1}^{\infty} \frac{(1-s+p_0/2)\mu_1^{n-2} - (1-s)(1-\alpha)^n \mu_1^{n-2}}{n!} - \sum_{n=1}^{\infty} \frac{(1-s+p_0/2)\mu_2^{n-2} - (1-s)(1-\alpha)^n \mu_2^{n-2}}{n!}}{\mu_2 - \mu_1} \\ &= (\mu_2 \mu_1) \sum_{n=1}^{\infty} \frac{(1-s+p_0/2 - (1-s)(1-\alpha)^n)(\mu_1^{n-2} - \mu_2^{n-2})}{n!(\mu_2 - \mu_1)} \\ &= p_0/2 + (1-s)\alpha + (\mu_2 \mu_1) \sum_{n=3}^{\infty} \frac{(1-s+p_0/2 - (1-s)(1-\alpha)^n)(\mu_1^{n-2} - \mu_2^{n-2})}{n!(\mu_2 - \mu_1)} \\ &= p_0/2 + (1-s)\alpha - (\mu_2 \mu_1) \sum_{n=3}^{\infty} \frac{1-s+p_0/2 - (1-s)(1-\alpha)^n}{n!} \left(\sum_{m=0}^{n-3} \mu_1^m \mu_2^{n-3-m} \right) \\ &= p_0/2 + (1-s)\alpha - \sum_{n=3}^{\infty} \frac{(1-s)(1 - (1-\alpha)^n) + p_0/2}{n!} \left(\sum_{m=0}^{n-3} \mu_1^{m+1} \mu_2^{n-2-m} \right). \end{aligned}$$

Here, $\frac{(1-s)(1-(1-\alpha)^n) + p_0/2}{n!}$ is always positive. Hence, $\hat{a}_\times^1(\mu_1, \mu_2)$ is monotone decreasing for μ_1 and μ_2 . \square

8.3. Case when true intensities are different from our intents

Next, we consider the case when we cannot perfectly identify the true intensities μ_1 and μ_2 . That is, we know that the true intensities μ_1 and μ_2 belong to certain intervals $[(1-\epsilon)\tilde{\mu}_1, (1+\epsilon)\tilde{\mu}_1]$ and $[(1-\epsilon)\tilde{\mu}_2, (1+\epsilon)\tilde{\mu}_2]$ with an constant $\epsilon > 0$, respectively. A similar analysis has been done by Wang et al.[13, 14]. However, our analysis based on the asymptotic key generation rate different from theirs.

In this case, we have to consider the worst case concerning the true intensities μ_1 and μ_2 in the intervals $[(1-\epsilon)\tilde{\mu}_1, (1+\epsilon)\tilde{\mu}_1]$ and $[(1-\epsilon)\tilde{\mu}_2, (1+\epsilon)\tilde{\mu}_2]$, respectively. Indeed, the smaller intensity pulse is generated by the combination of the stronger pulse and beam splitter. If the beam splitter is well installed, the error only comes from the error of the stronger pulse source. In this assumption, the error ratio ϵ does not depend on the intensity. Hence, when we observe the rates $p_{1,\times}$, $p_{2,\times}$, $s_{1,\times}$, and $s_{2,\times}$, the key generation rates per pulse with the coincidence basis are

$$\begin{aligned} R_1(p_{1,\times}, p_{2,\times}, s_{1,\times}, s_{2,\times}, \tilde{\mu}_1, \tilde{\mu}_2) \\ := \min \mu_1 e^{-\mu_1} (\hat{a}_\times^1 + \hat{b}_\times^1) (1 - h(\frac{\hat{b}_\times^1}{\hat{a}_\times^1 + \hat{b}_\times^1})) + e^{-\mu_1} p_0 - p_{1,\times} \eta h(\frac{s_{1,\times}}{p_{1,\times}}) \\ R_2(p_{1,\times}, p_{2,\times}, s_{1,\times}, s_{2,\times}, \tilde{\mu}_1, \tilde{\mu}_2) \\ := \min \mu_2 e^{-\mu_2} (\hat{a}_\times^1 + \hat{b}_\times^1) (1 - h(\frac{\hat{b}_\times^1}{\hat{a}_\times^1 + \hat{b}_\times^1})) + e^{-\mu_2} p_0 - p_{2,\times} \eta h(\frac{s_{2,\times}}{p_{2,\times}}), \end{aligned}$$

where the above minimums are taken with $\hat{a}_\times^1 = \min\{\hat{a}_\times^1(p_{1,\times} - s_{1,\times}, p_{2,\times} - s_{2,\times}, \mu_1, \mu_2), 0\}$ and $\hat{b}_\times^1 = \hat{b}_\times^1(s_{1,\times}, \mu_1)$ under the constraint $\mu_1 \in [(1-\epsilon)\tilde{\mu}_1, (1+\epsilon)\tilde{\mu}_1]$, $\mu_2 \in [(1-\epsilon)\tilde{\mu}_2, (1+\epsilon)\tilde{\mu}_2]$.

Now, we treat the typical case when true intensities are $\tilde{\mu}_1$ and $\tilde{\mu}_2$ and there is no eavesdropper. Then, the observed rates $p_{i,\times}$ and $s_{i,\times}$ are given by the true intensities $\tilde{\mu}_1$ and $\tilde{\mu}_2$ in the following way.

$$p_{i,\times} = 1 - e^{-\alpha\tilde{\mu}_i} + p_0, \quad s_{i,\times} = s(1 - e^{-\alpha\tilde{\mu}_i}) + \frac{p_0}{2}, \quad (115)$$

which implies $p_{i,\times} - s_{i,\times} = (1-s)(1 - e^{-\alpha\tilde{\mu}_i}) + \frac{p_0}{2}$. In order to analyze the key generation rates per pulse, we introduce additional notations.

$$\begin{aligned} \bar{R}_1(\mu_1, \mu_2, \tilde{\mu}_1, \tilde{\mu}_2) &:= \mu_1 e^{-\mu_1} (\hat{a}_\times^1 + \hat{b}_\times^1) (1 - h(\frac{\hat{b}_\times^1}{\hat{a}_\times^1 + \hat{b}_\times^1})) + e^{-\mu_1} p_0 - p_{1,\times} \eta h(\frac{s_{1,\times}}{p_{1,\times}}), \\ \bar{R}_2(\mu_1, \mu_2, \tilde{\mu}_1, \tilde{\mu}_2) &:= \mu_2 e^{-\mu_2} (\hat{a}_\times^1 + \hat{b}_\times^1) (1 - h(\frac{\hat{b}_\times^1}{\hat{a}_\times^1 + \hat{b}_\times^1})) + e^{-\mu_2} p_0 - p_{2,\times} \eta h(\frac{s_{2,\times}}{p_{2,\times}}), \end{aligned}$$

where $\hat{a}_\times^1 = \min\{\hat{a}_\times^1((1-s)(1 - e^{-\alpha\tilde{\mu}_1}) + \frac{p_0}{2}, (1-s)(1 - e^{-\alpha\tilde{\mu}_2}) + \frac{p_0}{2}, \mu_1, \mu_2), 0\}$ and $\hat{b}_\times^1 = \hat{b}_\times^1(s(1 - e^{-\alpha\tilde{\mu}_1}) + \frac{p_0}{2}, \mu_1)$. In this case, the key generation rates per pulse are given by

$$\begin{aligned} R_1(\tilde{\mu}_1, \tilde{\mu}_2) &:= \min_{\mu_1 \in [(1-\epsilon)\tilde{\mu}_1, (1+\epsilon)\tilde{\mu}_1], \mu_2 \in [(1-\epsilon)\tilde{\mu}_2, (1+\epsilon)\tilde{\mu}_2]} \bar{R}_1(\mu_1, \mu_2, \tilde{\mu}_1, \tilde{\mu}_2) \\ R_2(\tilde{\mu}_1, \tilde{\mu}_2) &:= \min_{\mu_1 \in [(1-\epsilon)\tilde{\mu}_1, (1+\epsilon)\tilde{\mu}_1], \mu_2 \in [(1-\epsilon)\tilde{\mu}_2, (1+\epsilon)\tilde{\mu}_2]} \bar{R}_2(\mu_1, \mu_2, \tilde{\mu}_1, \tilde{\mu}_2) \end{aligned}$$

Indeed, it is quite difficult to find the values $\mu_1 \in [(1 - \epsilon)\tilde{\mu}_1, (1 + \epsilon)\tilde{\mu}_1], \mu_2 \in [(1 - \epsilon)\tilde{\mu}_2, (1 + \epsilon)\tilde{\mu}_2]$ realizing the minimums $R_1(\tilde{\mu}_1, \tilde{\mu}_2)$ and $R_2(\tilde{\mu}_1, \tilde{\mu}_2)$. However, our numerical demonstration suggests that $\mu_1 = (1 + \epsilon)\tilde{\mu}_1$ and $\mu_2 = (1 - \epsilon)\tilde{\mu}_2$ give the minimums when $\epsilon > 0$ is sufficiently small and $(1 + \epsilon)\tilde{\mu}_1 < (1 - \epsilon)\tilde{\mu}_2$. Indeed, as is shown in Theorem 4, $\mu_1 = (1 + \epsilon)\tilde{\mu}_1$ gives the minimum $R_2(\tilde{\mu}_1, \tilde{\mu}_2)$ under the limit $\tilde{\mu}_1 \rightarrow 0$.

Theorem 4 *When a fixed intensity $\tilde{\mu}_2$ satisfies that*

$$\tilde{\mu}_2(1 + \epsilon) \leq 1 - \frac{p_0}{\left(\frac{\alpha}{1+\epsilon} + p_0\right)(1 - h(\frac{s\frac{\alpha}{1+\epsilon} + \frac{p_0}{2}}{\frac{\alpha}{1+\epsilon} + p_0})}, \quad (116)$$

we obtain

$$\begin{aligned} \sup_{\tilde{\mu}_1: 0 < \tilde{\mu}_1 < \tilde{\mu}_2} R_2(\tilde{\mu}_1, \tilde{\mu}_2) &= \lim_{\tilde{\mu}_1 \rightarrow 0} R_2(\tilde{\mu}_1, \tilde{\mu}_2) \\ &= \min_{\mu_2 \in [(1-\epsilon)\tilde{\mu}_2, (1+\epsilon)\tilde{\mu}_2]} \lim_{\tilde{\mu}_1 \rightarrow 0} \bar{R}_2((1 + \epsilon)\tilde{\mu}_1, \mu_2, \tilde{\mu}_1, \tilde{\mu}_2) \\ &= (1 + \epsilon)\tilde{\mu}_2 e^{-(1+\epsilon)\tilde{\mu}_2} \left(\frac{\alpha}{1 + \epsilon} + p_0 \right) (1 - h(\frac{s\frac{\alpha}{1+\epsilon} + \frac{p_0}{2}}{\frac{\alpha}{1+\epsilon} + p_0})) + e^{-(1+\epsilon)\tilde{\mu}_2} p_0 - p_{2,\times} \eta h(\frac{s_{2,\times}}{p_{2,\times}}), \end{aligned}$$

where $p_{2,\times}$ and $s_{2,\times}$ are given in (115).

For a proof of this theorem, we prepare the following lemmas.

Lemma 1 *When $\hat{a}_\times^1 = \hat{a}_\times^1((1 - s)(1 - e^{-\alpha\tilde{\mu}_1}) + \frac{p_0}{2}, (1 - s)(1 - e^{-\alpha\tilde{\mu}_2}) + \frac{p_0}{2}, \mu_1, \mu_2)$ and $\hat{b}_\times^1 = \hat{b}_\times^1(s(1 - e^{-\alpha\tilde{\mu}_1}) + \frac{p_0}{2}, \mu_1)$, \hat{a}_\times^1 is monotone increasing concerning $\tilde{\mu}_1$ and monotone decreasing concerning $\tilde{\mu}_2$, and $\frac{\hat{b}_\times^1}{\hat{a}_\times^1 + \hat{b}_\times^1}$ is monotone decreasing concerning $\tilde{\mu}_1$. Further, $(\hat{a}_\times^1 + \hat{b}_\times^1)(1 - h(\frac{\hat{b}_\times^1}{\hat{a}_\times^1 + \hat{b}_\times^1}))$ is monotone increasing concerning $\tilde{\mu}_1$ and monotone decreasing concerning $\tilde{\mu}_2$.*

Proof. The desired properties for \hat{a}_\times^1 follow from the fact that $(1 - s)(1 - e^{-\alpha\tilde{\mu}_i})$ is monotone increasing concerning $\tilde{\mu}_i$. Next, we consider $\frac{\hat{b}_\times^1}{\hat{a}_\times^1 + \hat{b}_\times^1}$. Since

$$\hat{b}_\times^1 = \frac{\mu_2(\mu_2 - \mu_1)e^{\mu_1}s(1 - e^{-\alpha\tilde{\mu}_1}) + \mu_2(\mu_2 - \mu_1)e^{\mu_1}(1 - e^{-\mu_1})p_0/2}{\mu_1\mu_2(\mu_2 - \mu_1)},$$

we have

$$\begin{aligned} \frac{\hat{b}_\times^1}{\hat{a}_\times^1 + \hat{b}_\times^1} &= \frac{\mu_2 e^{\mu_1}(\mu_2 - \mu_1)s(1 - e^{-\alpha\tilde{\mu}_1}) + \mu_2(\mu_2 - \mu_1)e^{\mu_1}(1 - e^{-\mu_1})p_0/2}{\mu_2 e^{\mu_1}(\mu_2 - \mu_1)s(1 - e^{-\alpha\tilde{\mu}_1}) + \mu_2 e^{\mu_1}(1 - e^{-\mu_1})(\mu_2 - \mu_1/2)p_0 - B_1} \\ &= \frac{B_2(1 - e^{-\alpha\tilde{\mu}_1}) + B_3}{B_4(1 - e^{-\alpha\tilde{\mu}_1}) + B_5}, \end{aligned}$$

where

$$B_1 := \mu_1^2 e^{\mu_2}((1 - s)(1 - e^{-\alpha\tilde{\mu}_2}) + (1 - e^{-\mu_2})p_0/2)$$

$$B_2 := \mu_2 e^{\mu_1}(\mu_2 - \mu_1)s$$

$$B_3 := \mu_2(\mu_2 - \mu_1)e^{\mu_1}(1 - e^{-\mu_1})p_0/2$$

$$B_4 := \mu_2 e^{\mu_1}(\mu_2 - \mu_1)s$$

$$B_5 := \mu_2 e^{\mu_1}(1 - e^{-\mu_1})(\mu_2 - \mu_1/2)p_0 - B_1.$$

Since $s < \frac{1}{2}$, we have $\frac{s}{(\mu_2 - \mu_1 s)} < \frac{1}{2(\mu_2 - \mu_1/2)}$, which implies

$$\begin{aligned} \frac{B_2}{B_4} &= \frac{\mu_2 e^{\mu_1} (\mu_2 - \mu_1) s}{\mu_2 e^{\mu_1} (\mu_2 - \mu_1 s)} = \frac{s(\mu_2 - \mu_1)}{(\mu_2 - \mu_1 s)} < \frac{\mu_2 - \mu_1}{2(\mu_2 - \mu_1/2)} \\ &= \frac{\mu_2 (\mu_2 - \mu_1) e^{\mu_1} (1 - e^{-\mu_1}) p_0 / 2}{\mu_2 e^{\mu_1} (1 - e^{-\mu_1}) (\mu_2 - \mu_1/2) p_0} < \frac{\mu_2 (\mu_2 - \mu_1) e^{\mu_1} (1 - e^{-\mu_1}) p_0 / 2}{\mu_2 e^{\mu_1} (1 - e^{-\mu_1}) (\mu_2 - \mu_1/2) p_0 - B_1} = \frac{B_3}{B_5}. \end{aligned}$$

Since $1 - e^{-\alpha \tilde{\mu}_1}$ is monotone increasing concerning $\tilde{\mu}_1$, $\frac{B_2(1 - e^{-\alpha \tilde{\mu}_1}) + B_3}{B_4(1 - e^{-\alpha \tilde{\mu}_1}) + B_5}$ is monotone decreasing concerning $\tilde{\mu}_1$. Then, we obtain the desired argument for $\frac{\hat{b}_x^1}{\hat{a}_x^1 + \hat{b}_x^1}$.

Finally, we use the function $f(a, b)$ defined in the proof of Theorem 3. Since $f(a, b)$ is monotone increasing concerning a and monotone decreasing concerning b , we obtain the desired argument for $(\hat{a}_x^1 + \hat{b}_x^1)(1 - h(\frac{\hat{b}_x^1}{\hat{a}_x^1 + \hat{b}_x^1}))$. \square

Lemma 2 When $\tilde{\mu}_2 \geq (1 + \epsilon)^{-1} \mu_2$,

$$\begin{aligned} &\sup_{\mu_1: 0 < \mu_1 < \mu_2} \hat{a}_x^1((1 - s)(1 - e^{-\alpha(1+\epsilon)^{-1}\mu_1}) + \frac{p_0}{2}, (1 - s)(1 - e^{-\alpha\tilde{\mu}_2}) + \frac{p_0}{2}, \mu_1, \mu_2) \\ &= \lim_{\mu_1 \rightarrow +0} \hat{a}_x^1((1 - s)(1 - e^{-\alpha(1+\epsilon)^{-1}\mu_1}) + \frac{p_0}{2}, (1 - s)(1 - e^{-\alpha\tilde{\mu}_2}) + \frac{p_0}{2}, \mu_1, \mu_2) \\ &= (1 - s)\alpha(1 + \epsilon)^{-1} + \frac{p_0}{2}, \end{aligned} \tag{117}$$

$$\begin{aligned} &\sup_{0 < \mu_1} \hat{b}_x^1(s(1 - e^{-\alpha(1+\epsilon)^{-1}\mu_1}) + \frac{p_0}{2}, \mu_1) \\ &= \lim_{\mu_1 \rightarrow +0} \hat{b}_x^1(s(1 - e^{-\alpha(1+\epsilon)^{-1}\mu_1}) + \frac{p_0}{2}, \mu_1) = s\alpha(1 + \epsilon)^{-1} + \frac{p_0}{2}. \end{aligned} \tag{118}$$

Proof. Since $\hat{b}_x^1(s(1 - e^{-\alpha(1+\epsilon)^{-1}\mu_1}) + \frac{p_0}{2}, \mu_1)$ is monotone decreasing concerning μ_1 , we obtain (118). Similarly, it is sufficient for (117) to show that $\hat{a}_x^1((1 - s)(1 - e^{-\alpha(1+\epsilon)^{-1}\mu_1}) + \frac{p_0}{2}, (1 - s)(1 - e^{-\alpha\tilde{\mu}_2}) + \frac{p_0}{2}, \mu_1, \mu_2)$ is monotone decreasing concerning μ_1 . Choosing $\tilde{\alpha} := \alpha(1 + \epsilon)^{-1}$, we obtain

$$\begin{aligned} &\hat{a}_x^1((1 - s)(1 - e^{-\alpha(1+\epsilon)^{-1}\mu_1}) + \frac{p_0}{2}, (1 - s)(1 - e^{-\alpha\tilde{\mu}_2}) + \frac{p_0}{2}, \mu_1, \mu_2) \\ &= \frac{\mu_2^2 e^{\mu_1} ((1 - s)(1 - e^{-\tilde{\alpha}\mu_1}) + p_0(1 - e^{-\mu_1})/2) - \mu_1^2 e^{\mu_2} ((1 - s)(1 - e^{-\alpha\mu_2}) + p_0(1 - e^{-\mu_2})/2)}{\mu_1 \mu_2 (\mu_2 - \mu_1)} \\ &= \frac{\mu_2^2 e^{\mu_1} ((1 - s)(1 - e^{-\tilde{\alpha}\mu_1}) + p_0(1 - e^{-\mu_1})/2) - \mu_1^2 e^{\mu_2} ((1 - s)(1 - e^{-\tilde{\alpha}\mu_2}) + p_0(1 - e^{-\mu_2})/2)}{\mu_1 \mu_2 (\mu_2 - \mu_1)} \\ &\quad - \frac{\mu_1 e^{\mu_2} ((1 - s)(e^{-\tilde{\alpha}\mu_2} - e^{-\alpha\mu_2}))}{\mu_2 (\mu_2 - \mu_1)} \\ &= p_0/2 + (1 - s)\tilde{\alpha} - \sum_{n=3}^{\infty} \frac{(1 - s)(1 - (1 - \alpha)^n) + p_0/2}{n!} \left(\sum_{m=0}^{n-3} \mu_1^{m+1} \mu_2^{n-2-m} \right) \\ &\quad - \frac{\mu_1 e^{\mu_2} ((1 - s)(e^{-\tilde{\alpha}\mu_2} - e^{-\alpha\mu_2}))}{\mu_2 (\mu_2 - \mu_1)}. \end{aligned}$$

Since $e^{-\tilde{\alpha}\mu_2} - e^{-\alpha\mu_2} \geq 0$, the final term is monotone decreasing concerning μ_1 . Other terms are also monotone decreasing concerning μ_1 . \square

Proof of Theorem 4.

For a fixed $\tilde{\mu}_2$ and μ_2 satisfying $(1 - \epsilon)\tilde{\mu}_2 < \mu_2 < (1 + \epsilon)\tilde{\mu}_2$, Lemmas 1 and 2 imply

$$\begin{aligned}
& \sup_{\tilde{\mu}_1: 0 < \tilde{\mu}_1 < \frac{1-\epsilon}{1+\epsilon}\tilde{\mu}_2} \min_{\mu_1 \in [(1-\epsilon)\tilde{\mu}_1, (1+\epsilon)\tilde{\mu}_1]} \bar{R}_2(\mu_1, \mu_2, \tilde{\mu}_1, \tilde{\mu}_2) \\
&= \sup_{\mu_1: 0 < \mu_1 < \mu_2} \min_{\tilde{\mu}_1: (1-\epsilon)\tilde{\mu}_1 < \mu_1 < (1+\epsilon)\tilde{\mu}_1} \bar{R}_2(\mu_1, \mu_2, \tilde{\mu}_1, \tilde{\mu}_2) \\
&= \sup_{\mu_1: 0 < \mu_1 < \mu_2} \bar{R}_2(\mu_1, \mu_2, (1 + \epsilon)^{-1}\mu_1, \tilde{\mu}_2) \\
&= \lim_{\mu_1 \rightarrow +0} \bar{R}_2(\mu_1, \mu_2, (1 + \epsilon)^{-1}\mu_1, \tilde{\mu}_2). \tag{119}
\end{aligned}$$

Hence, we obtain

$$\begin{aligned}
& \sup_{\tilde{\mu}_1: 0 < \tilde{\mu}_1 < \tilde{\mu}_2} \min_{\mu_1 \in [(1-\epsilon)\tilde{\mu}_1, (1+\epsilon)\tilde{\mu}_1]} \bar{R}_2(\mu_1, \mu_2, \tilde{\mu}_1, \tilde{\mu}_2) \\
&= \lim_{\tilde{\mu}_1 \rightarrow +0} \min_{\mu_1 \in [(1-\epsilon)\tilde{\mu}_1, (1+\epsilon)\tilde{\mu}_1]} \bar{R}_2(\mu_1, \mu_2, \tilde{\mu}_1, \tilde{\mu}_2) \\
&= \lim_{\tilde{\mu}_1 \rightarrow +0} \bar{R}_2((1 + \epsilon)\tilde{\mu}_1, \mu_2, \tilde{\mu}_1, \tilde{\mu}_2). \tag{120}
\end{aligned}$$

Since the convergence (119) is uniform concerning μ_2 and $\tilde{\mu}_2$, the convergence (120) is uniform concerning μ_2 and $\tilde{\mu}_2$. Hence, we obtain

$$\begin{aligned}
& \min_{\mu_2 \in [(1-\epsilon)\tilde{\mu}_2, (1+\epsilon)\tilde{\mu}_2]} \lim_{\tilde{\mu}_1 \rightarrow +0} \min_{\mu_1 \in [(1-\epsilon)\tilde{\mu}_1, (1+\epsilon)\tilde{\mu}_1]} \bar{R}_2(\mu_1, \mu_2, \tilde{\mu}_1, \tilde{\mu}_2) \\
&= \lim_{\tilde{\mu}_1 \rightarrow +0} \min_{\mu_2 \in [(1-\epsilon)\tilde{\mu}_2, (1+\epsilon)\tilde{\mu}_2]} \min_{\mu_1 \in [(1-\epsilon)\tilde{\mu}_1, (1+\epsilon)\tilde{\mu}_1]} \bar{R}_2(\mu_1, \mu_2, \tilde{\mu}_1, \tilde{\mu}_2). \tag{121}
\end{aligned}$$

On the other hand,

$$\begin{aligned}
& \min_{\mu_2 \in [(1-\epsilon)\tilde{\mu}_2, (1+\epsilon)\tilde{\mu}_2]} \lim_{\tilde{\mu}_1 \rightarrow +0} \min_{\mu_1 \in [(1-\epsilon)\tilde{\mu}_1, (1+\epsilon)\tilde{\mu}_1]} \bar{R}_2(\mu_1, \mu_2, \tilde{\mu}_1, \tilde{\mu}_2) \\
&= \min_{\mu_2 \in [(1-\epsilon)\tilde{\mu}_2, (1+\epsilon)\tilde{\mu}_2]} \sup_{\tilde{\mu}_1: 0 < \tilde{\mu}_1 < \frac{1-\epsilon}{1+\epsilon}\tilde{\mu}_2} \min_{\mu_1 \in [(1-\epsilon)\tilde{\mu}_1, (1+\epsilon)\tilde{\mu}_1]} \bar{R}_2(\mu_1, \mu_2, \tilde{\mu}_1, \tilde{\mu}_2) \\
&\geq \sup_{\tilde{\mu}_1: 0 < \tilde{\mu}_1 < \frac{1-\epsilon}{1+\epsilon}\tilde{\mu}_2} \min_{\mu_2 \in [(1-\epsilon)\tilde{\mu}_2, (1+\epsilon)\tilde{\mu}_2]} \min_{\mu_1 \in [(1-\epsilon)\tilde{\mu}_1, (1+\epsilon)\tilde{\mu}_1]} \bar{R}_2(\mu_1, \mu_2, \tilde{\mu}_1, \tilde{\mu}_2) \\
&\geq \lim_{\tilde{\mu}_1 \rightarrow +0} \min_{\mu_2 \in [(1-\epsilon)\tilde{\mu}_2, (1+\epsilon)\tilde{\mu}_2]} \min_{\mu_1 \in [(1-\epsilon)\tilde{\mu}_1, (1+\epsilon)\tilde{\mu}_1]} \bar{R}_2(\mu_1, \mu_2, \tilde{\mu}_1, \tilde{\mu}_2). \tag{122}
\end{aligned}$$

Combining (121) and (122), we obtain

$$\begin{aligned}
& \sup_{\tilde{\mu}_1: 0 < \tilde{\mu}_1 < \frac{1-\epsilon}{1+\epsilon}\tilde{\mu}_2} \min_{\mu_2 \in [(1-\epsilon)\tilde{\mu}_2, (1+\epsilon)\tilde{\mu}_2]} \min_{\mu_1 \in [(1-\epsilon)\tilde{\mu}_1, (1+\epsilon)\tilde{\mu}_1]} \bar{R}_2(\mu_1, \mu_2, \tilde{\mu}_1, \tilde{\mu}_2) \\
&= \min_{\mu_2 \in [(1-\epsilon)\tilde{\mu}_2, (1+\epsilon)\tilde{\mu}_2]} \lim_{\tilde{\mu}_1 \rightarrow +0} \min_{\mu_1 \in [(1-\epsilon)\tilde{\mu}_1, (1+\epsilon)\tilde{\mu}_1]} \bar{R}_2(\mu_1, \mu_2, \tilde{\mu}_1, \tilde{\mu}_2) \\
&= \min_{\mu_2 \in [(1-\epsilon)\tilde{\mu}_2, (1+\epsilon)\tilde{\mu}_2]} \lim_{\tilde{\mu}_1 \rightarrow +0} \bar{R}_2((1 + \epsilon)\tilde{\mu}_1, \mu_2, \tilde{\mu}_1, \tilde{\mu}_2) \\
&= \min_{\mu_2 \in [(1-\epsilon)\tilde{\mu}_2, (1+\epsilon)\tilde{\mu}_2]} \lim_{\mu_1 \rightarrow +0} \bar{R}_2(\mu_1, \mu_2, (1 + \epsilon)^{-1}\mu_1, \tilde{\mu}_2) \\
&= \min_{\mu_2 \in [(1-\epsilon)\tilde{\mu}_2, (1+\epsilon)\tilde{\mu}_2]} \mu_2 e^{-\mu_2} \left(\frac{\alpha}{1 + \epsilon} + p_0 \right) \left(1 - h\left(\frac{s \frac{\alpha}{1+\epsilon} + \frac{p_0}{2}}{\frac{\alpha}{1+\epsilon} + p_0} \right) \right) + e^{-\mu_2} p_0 - p_{2,\times} \eta h\left(\frac{s_{2,\times}}{p_{2,\times}} \right) \\
&= (1 + \epsilon)\tilde{\mu}_2 e^{-(1+\epsilon)\tilde{\mu}_2} \left(\frac{\alpha}{1 + \epsilon} + p_0 \right) \left(1 - h\left(\frac{s \frac{\alpha}{1+\epsilon} + \frac{p_0}{2}}{\frac{\alpha}{1+\epsilon} + p_0} \right) \right) + e^{-(1+\epsilon)\tilde{\mu}_2} p_0 - p_{2,\times} \eta h\left(\frac{s_{2,\times}}{p_{2,\times}} \right),
\end{aligned}$$

where the final equation follows from (116) and the following fact. The function $x \mapsto xe^{-x} + ae^{-x}$ is monotone increasing when $x \leq 1 - a$ for $a > 0$. \square

The rates $R_1(\tilde{\mu}_1, \tilde{\mu}_2)$ and $R_2(\tilde{\mu}_1, \tilde{\mu}_2)$ are numerically calculated with $s = 0.03$, $p_0 = 4.0 \times 10^{-7}$, $\alpha = 1.0 \times 10^{-3}$ as Figures 5, 6, 8, 7, and 9. Indeed, when $\epsilon = 0$ and $\tilde{\mu}_1$ is fixed, the rate R_1 can be maximized with the limit $\tilde{\mu}_2 \rightarrow \tilde{\mu}_1$. However, as is shown in Figure 5, when there exists error, the rate R_1 goes to zero with the limit $\tilde{\mu}_2 \rightarrow \tilde{\mu}_1$. As is shown in Figures 5 and 8, Both rates R_1 and R_2 become zero with the limit $\tilde{\mu}_2 \rightarrow \tilde{\mu}_1$. This is because \hat{a}_x^1 goes to zero since two intensities cannot be distinguished when $\tilde{\mu}_2$ is close to $\tilde{\mu}_1$. Figure 6 and Table 1 give the optimal intensities $\tilde{\mu}_1$ and $\tilde{\mu}_2$ for R_1 .

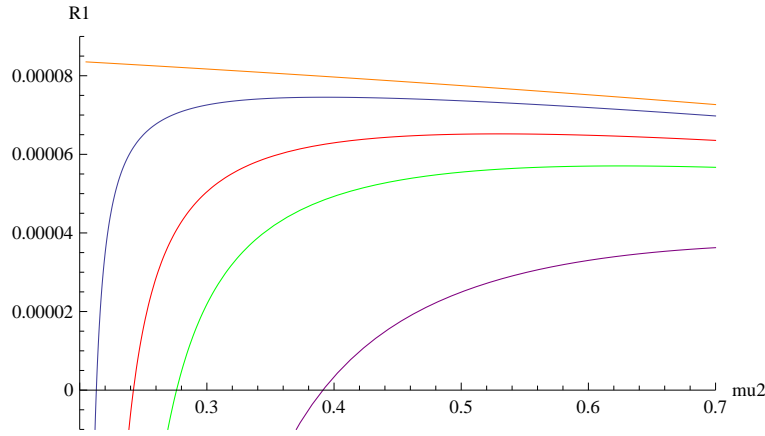


Figure 5. Key generation rate R_1 with $\tilde{\mu}_1 = 0.1$. The horizontal axis is $\tilde{\mu}_2$. The error parameter ϵ is chosen to be 0% (orange) 1% (blue), 3% (red), 5% (green), and 10% (purple).

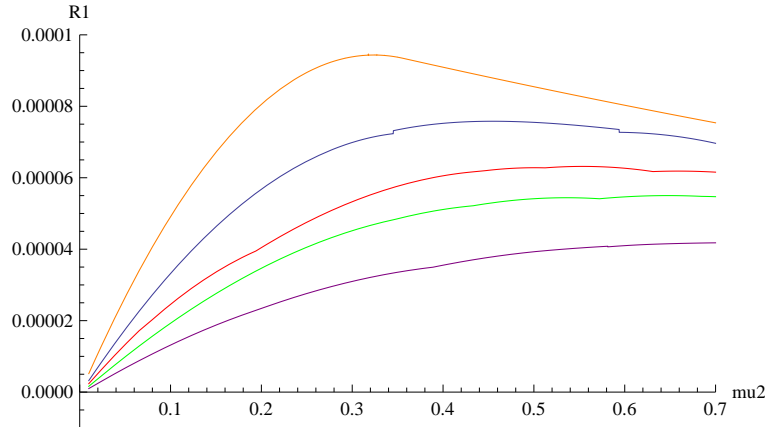


Figure 6. Key generation rate $\max_{0 < \tilde{\mu}_1 < \tilde{\mu}_2} R_1(\tilde{\mu}_1, \tilde{\mu}_2)$. The horizontal axis is $\tilde{\mu}_2$. The error parameter ϵ is chosen to be 0% (orange) 1% (blue), 3% (red), 5% (green), and 10% (purple).

When $\tilde{\mu}_1$ is close to zero, the the difference $\tilde{\mu}_1 - \mu_1$ becomes small in proportion to $\tilde{\mu}_1$. Then, as is shown in Theorem 4, the limit $\tilde{\mu}_1 \rightarrow 0$ gives the optimal asymptotic rate R_2 for a fixed $\tilde{\mu}_2$. Figure 7 and Table 2 give the optimal signal intensity $\tilde{\mu}_2$ for R_2 . However, in a real setting, it is quite difficult to control too small intensity. Hence, we consider the optimization of $\tilde{\mu}_2$ for R_2 when $\mu_1 = 0.1$, as is shown Figure 8 and Table 3.

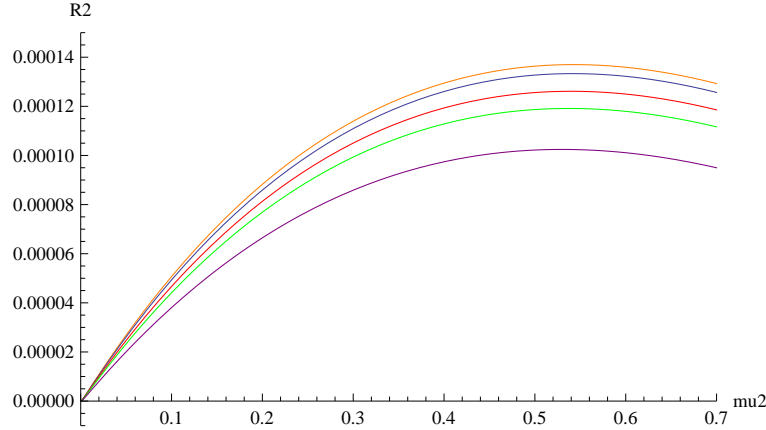


Figure 7. Key generation rate $\max_{0 < \tilde{\mu}_1 < \tilde{\mu}_2} R_2(\tilde{\mu}_1, \tilde{\mu}_2)$. The horizontal axis is $\tilde{\mu}_2$. The error parameter ϵ is chosen to be 0% (orange) 1% (blue), 3% (red), 5% (green), and 10% (purple).

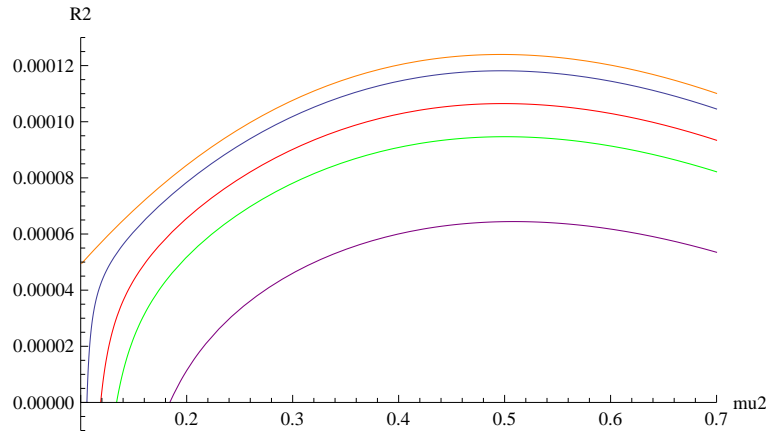


Figure 8. Key generation rate R_2 with $\tilde{\mu}_1 = 0.2$. The horizontal axis is $\tilde{\mu}_2$. The error parameter ϵ is chosen to be 0% (orange) 1% (blue), 3% (red), 5% (green), and 10% (purple).

Note that the rate R_2 assumes the asymptotic limit concerning the number of coding length. If we take into account finiteness of the number of coding length, we have to consider the statistical fluctuations of \vec{N} and \mathbf{M} . This problem will be discussed in Section 11.

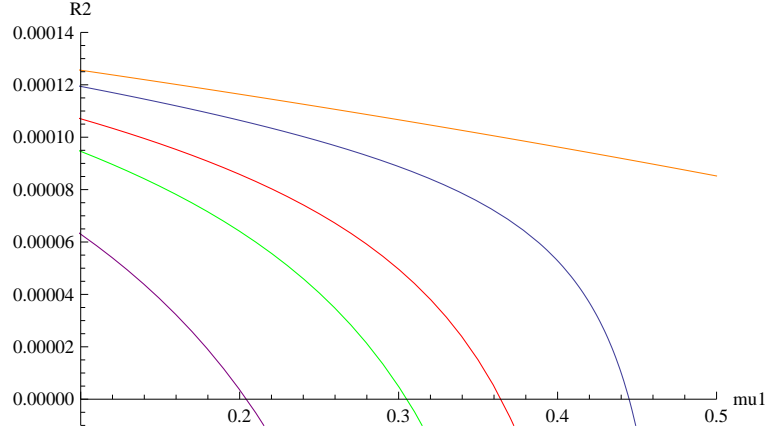


Figure 9. Key generation rate R_2 with $\tilde{\mu}_2 = 0.5$. The horizontal axis is $\tilde{\mu}_1$. The error parameter ϵ is chosen to be 0% (orange) 1% (blue), 3% (red), 5% (green), and 10% (purple).

Table 1. Optimal intensities for R_1 and optimal R_1

ϵ	$\tilde{\mu}_1$	$\tilde{\mu}_2$	R_1
0%	0.323867	0.323868	0.000094349
1%	0.260721	0.454851	0.0000758128
3%	0.364856	0.381122	0.00139591
5%	0.190192	0.613888	0.0000551087
10%	0.145002	0.7413	0.0000418435

Table 2. Optimal intensities for R_2 and optimal R_2

ϵ	$\tilde{\mu}_1$	$\tilde{\mu}_2$	R_1
0%	0	0.542685	0.000136993
1%	0	0.541796	0.000133319
3%	0	0.539778	0.000126128
5%	0	0.537399	0.000119136
10%	0	0.529685	0.000102456

Table 3. Optimal intensities for R_2 and optimal R_2 when $\tilde{\mu}_1 = 0.1$

ϵ	$\tilde{\mu}_2$	R_1
0%	0.49672	0.000123965
1%	0.49719	0.000118152
3%	0.498509	0.000106466
5%	0.500422	0.0000946697
10%	0.482252	0.0000641944

9. Derivation of upper bound $\hat{\phi}_4$ of leaked information

9.1. Condition for μ_1 , μ_2 , and \mathbf{M}

In this section, we define the upper bound $\hat{\phi}_4(\mathbf{M})$ satisfying (42) as a function of the measured value \mathbf{M} that does not depend on the breakdown $\vec{\mathbf{N}}$. In the following, we

treat the case when the signal intensity is μ_1 and the decoy intensity is μ_2 . However, the discussions in this section except for Subsection 9.2 are still valid with replacing N , N^0 , and N^1 by N' , $N^{0'}$, and $N^{1'}$ even when the signal intensity is μ_2 and the decoy intensity is μ_1 .

In this subsection, for this purpose, we introduce the following condition for μ_1 , μ_2 , and \mathbf{M} .

Condition 6 For any $\vec{N} \in \Omega_1$, all of the following values are positive.

$$\begin{aligned} A_1^0 &:= \hat{M}_2 - \hat{M}_0(N_2^0 + N_2^2)/2N_0 - N_2^1 \frac{N_2^2(\hat{M}_1 - \hat{M}_0 N_1^0/2N_0) - N_1^2(\hat{M}_2 - \hat{M}_0 N_2^0/2N_0)}{N_1^1 N_2^2 - N_2^1 N_1^2}, \\ A_1^1 &:= \hat{M}_2 - \hat{M}_0 N_2^0/2N_0 - (N_2^2 + N_2^1) \frac{N_2^2(\hat{M}_1 - \hat{M}_0 N_1^0/2N_0) - N_1^2(\hat{M}_2 - \hat{M}_0 N_2^0/2N_0)}{N_1^1 N_2^2 - N_2^1 N_1^2}, \\ A_2^1 &:= \frac{N_1^2(N_2^2(\hat{M}_1 - \frac{\hat{M}_0}{2N_0} N_1^0) - N_1^2(\hat{M}_2 - \frac{\hat{M}_0}{2N_0} N_2^0))}{N_1^1 N_2^2 - N_2^1 N_1^2}, \\ A_2^2 &:= \hat{M}_1 - \frac{\hat{M}_0}{2N_0} N_1^0 - \frac{N_1^1(N_2^2(\hat{M}_1 - \frac{\hat{M}_0}{2N_0} N_1^0) - N_1^2(\hat{M}_2 - \frac{\hat{M}_0}{2N_0} N_2^0))}{N_1^1 N_2^2 - N_2^1 N_1^2}, \\ B_1^1 &:= \hat{M}_3 - \hat{M}_0 N_1^0/2N_0. \end{aligned}$$

Substituting $\bar{\mathbf{M}}$ into $\hat{\mathbf{M}}$, and applying the relations (77), (78), and (79), we can calculate the above values as

$$A_1^0 = (\hat{a}_\times^1 - \hat{q}^0/2)N_2^2, \quad A_1^1 = (\hat{a}_\times^2 - \hat{a}_\times^1)N_2^2, \quad A_2^1 = \hat{a}_\times^1 N_1^2, \quad A_2^2 = \hat{a}_\times^2 N_2^2, \quad B_1^1 = \hat{b}_\times^1 N_1^1.$$

Now, we show that these values take positive values, naturally. As is shown in the beginning of Section 6, when all of \vec{N} are close to their expectations, and there is no eavesdrop, i.e., the condition (104) holds, $\hat{a}_\times^2 - \hat{q}^0/2$ and $\hat{a}_\times^2 - \hat{a}_\times^1$ are positive. Thus, Condition 6 holds under the condition (104). Since the condition (104) is a natural assumption, Condition 6 can be regarded as a natural assumption. Hence, we need to choose the parameters μ_1 , μ_2 , N , N_0 , N_1 , N_2 so that Condition 6 holds with high probability without an eavesdropper.

Here, we need to pay attention to the difference between Conditions 6 and 5. Condition 5 is an assumption for μ_1 , μ_2 , N , N_0 , N_1 , and N_2 . On the other hand, Condition 6 is an assumption for the measured values \mathbf{M} and μ_1 , μ_2 because the estimates $\hat{\mathbf{M}}$ are determined from the measured values \mathbf{M} via the relations (91), (92), (93), and (94).

When there exists an eavesdropper, even if we choose μ_1 , μ_2 , N , N_0 , N_1 , N_2 suitably, the eavesdropper might control the channel parameters \bar{q}^0 , $\bar{\mathbf{a}}$ and $\bar{\mathbf{b}}$ so that Condition 6 does not hold. Hence, we need to prepare a method to smoothly decide whether Condition 6 holds.

9.2. Derivation of \mathbf{N}

Next, we derive the pair (\bar{N}^0, \bar{N}^1) to give $\hat{\phi}_1$ when \hat{q}^0 , \hat{a}_\times^1 , and \hat{b}_\times^1 are given. Only $-\hat{J}^0(\bar{q}^0, N^0)$ depends on N^0 in the right hand side of (56), and is monotone decreasing

concerning N^0 . Hence, $\hat{\phi}_1(\mathbf{N}, \bar{q}^0, \bar{q}^1, \bar{r}_\times^1)$ is monotone decreasing concerning N^0 .

On the other hand, only $-\hat{J}^1(\bar{q}^1, N^1)(1 - h(\min\{\hat{r}_\times^1(\bar{q}^1, \bar{b}_\times^1, N^1), 1/2\}))$ depends on N^1 the right hand side of (56). The parameters \hat{r}_\times^1 and \hat{r}_\times^1 are monotone decreasing concerning N^1 and \hat{J}^1 is monotone increasing concerning N^1 . Since $(h(\hat{r}_\times^1) - 1)$ is negative and h is a monotone increasing function with an input less than $1/2$, $\hat{J}^1(h(\hat{r}_\times^1) - 1)$ is monotone decreasing concerning N^1 .

Since N^0 and N^1 obey the binomial distributions $\text{Bin}(N, e^{-\mu_1})$ and $\text{Bin}(N, \mu_1 e^{-\mu_1})$, respectively, when the signal intensity is μ_1 and the decoy intensity is μ_2 , we give the estimate $\hat{\mathbf{N}} = (\hat{N}^0, \hat{N}^1)$ as

$$\hat{N}^0 := Y_1^-(N, e^{-\mu_1}, 2^{-2\beta-8}), \quad \hat{N}^1 := Y_1^-(N, \mu_1 e^{-\mu_1}, 2^{-2\beta-8}).$$

Similarly, when the signal intensity is μ_2 and the decoy intensity is μ_1 , we give the estimate $\hat{\mathbf{N}} = (\hat{N}^0, \hat{N}^1)$ as

$$\hat{N}^0 := Y_1^-(N', e^{-\mu_2}, 2^{-2\beta-8}), \quad \hat{N}^1 := Y_1^-(N', \mu_2 e^{-\mu_2}, 2^{-2\beta-8}).$$

9.3. Derivation of \mathbf{N}_1 and \mathbf{N}_2

Next, in order to derive the estimates of \mathbf{N}_1 and \mathbf{N}_2 giving an upper bound of $\hat{\phi}_2(\mathbf{M})$, we calculate the partial derivatives of $\hat{\phi}_1$ concerning \mathbf{N}_1 and \mathbf{N}_2 . For this purpose, we calculate the partial derivatives of $\hat{a}_\times^1(\bar{\mathbf{M}}, \bar{\mathbf{N}})$ concerning $N_1^0, N_1^1, N_2^0, N_2^1, N_2^2$ as follows.

$$\begin{aligned} & \frac{\partial \hat{a}_\times^1(\hat{\mathbf{M}}_0, \mathbf{N}_1, \mathbf{N}_2)}{\partial N_1^0} \\ &= \frac{\hat{M}_2 - \hat{M}_0(N_2^0 + N_2^2)/2N_0}{N_1^1 N_2^2 - N_2^1 N_1^2} - N_2^1 \frac{N_2^2(\hat{M}_1 - \hat{M}_0 N_1^0/2N_0) - N_1^2(\hat{M}_2 - \hat{M}_0 N_2^0/2N_0)}{(N_1^1 N_2^2 - N_2^1 N_1^2)^2}, \\ & \frac{\partial \hat{a}_\times^1(\hat{\mathbf{M}}, \mathbf{N}_1, \mathbf{N}_2)}{\partial N_1^1} \\ &= \frac{\hat{M}_2 - \hat{M}_0 N_2^0/2N_0}{N_1^1 N_2^2 - N_2^1 N_1^2} - (N_2^2 + N_2^1) \frac{N_2^2(\hat{M}_1 - \hat{M}_0 N_1^0/2N_0) - N_1^2(\hat{M}_2 - \hat{M}_0 N_2^0/2N_0)}{(N_1^1 N_2^2 - N_2^1 N_1^2)^2}, \\ & \frac{\partial \hat{a}_\times^1(\hat{\mathbf{M}}, \mathbf{N}_1, \mathbf{N}_2)}{\partial N_2^0} = \frac{N_1^2 \hat{M}_0/2N_0}{N_1^1 N_2^2 - N_2^1 N_1^2}, \\ & \frac{\partial \hat{a}_\times^1(\hat{\mathbf{M}}, \mathbf{N}_1, \mathbf{N}_2)}{\partial N_2^1} = \frac{N_1^2(N_2^2(\hat{M}_1 - \frac{\hat{M}_0}{2N_0} N_1^0) - N_1^2(\hat{M}_2 - \frac{\hat{M}_0}{2N_0} N_2^0))}{(N_1^1 N_2^2 - N_2^1 N_1^2)^2}, \\ & \frac{\partial \hat{a}_\times^1(\hat{\mathbf{M}}_0, \mathbf{N}_1, \mathbf{N}_2)}{\partial N_2^2} = \frac{\hat{M}_1 - \frac{\hat{M}_0}{2N_0} N_1^0}{N_1^1 N_2^2 - N_2^1 N_1^2} - \frac{N_1^1(N_2^2(\hat{M}_1 - \frac{\hat{M}_0}{2N_0} N_1^0) - N_1^2(\hat{M}_2 - \frac{\hat{M}_0}{2N_0} N_2^0))}{(N_1^1 N_2^2 - N_2^1 N_1^2)^2}. \end{aligned}$$

Here, we should we remark that $N_1^2 = N_1 - N_1^0 - N_1^1$. That is, the variable N_1^2 is a dependent variable. Due to Condition 6, all of the above values are positive.

Next, under Condition 6, we calculate the partial derivatives of $\hat{b}_\times^1(\bar{\mathbf{M}}, \bar{\mathbf{N}})$ concerning $N_1^0, N_1^1, N_2^0, N_2^1, N_2^2$ as follows.

$$\frac{\partial \hat{b}_\times^1(\hat{\mathbf{M}}_0, \mathbf{N}_1, \mathbf{N}_2)}{\partial N_1^0} = -\frac{\hat{M}_0}{2N_1^1 N_0} < 0$$

$$\frac{\partial \hat{b}_\times^1(\hat{\mathbf{M}}, \mathbf{N}_1, \mathbf{N}_2)}{\partial N_1^1} = -\frac{\hat{M}_3 - \hat{M}_0 N_1^0 / 2N_0}{(N_1^1)^2} < 0$$

$$\frac{\partial \hat{b}_\times^1(\hat{\mathbf{M}}, \mathbf{N}_1, \mathbf{N}_2)}{\partial N_2^0} = \frac{\partial \hat{b}_\times^1(\hat{\mathbf{M}}, \mathbf{N}_1, \mathbf{N}_2)}{\partial N_2^1} = \frac{\partial \hat{b}_\times^1(\hat{\mathbf{M}}, \mathbf{N}_1, \mathbf{N}_2)}{\partial N_2^2} = 0.$$

Since $\frac{\partial \hat{\phi}_1}{\partial \hat{a}_\times^1} < 0$ and $\frac{\partial \hat{\phi}_1}{\partial \hat{b}_\times^1} > 0$, due to (63) and (64), any element $\vec{\mathbf{N}} \in \Omega_1$ satisfies

$$\frac{\partial \hat{\phi}_1}{\partial N_1^0} < 0, \quad \frac{\partial \hat{\phi}_1}{\partial N_1^1} < 0, \quad \frac{\partial \hat{\phi}_1}{\partial N_2^0} < 0, \quad \frac{\partial \hat{\phi}_1}{\partial N_2^1} < 0, \quad \frac{\partial \hat{\phi}_1}{\partial N_2^2} < 0.$$

Thus, since $N_1^0, N_1^1, N_2^0, N_2^1$, and N_2^2 obey the binomial distribution, we decide $\hat{\mathbf{N}}_1 = (\hat{N}_1^0, \hat{N}_1^1)$ and $\hat{\mathbf{N}}_2 = (\hat{N}_2^0, \hat{N}_2^1, \hat{N}_2^2)$ in the following way:

$$\begin{aligned} \hat{N}_1^0 &:= Y_1^-(N_1, e^{-\mu_1}, 2^{-2\beta-8}) \\ \hat{N}_1^1 &:= Y_1^-(N_1, \mu_1 e^{-\mu_1}, 2^{-2\beta-8}) \\ \hat{N}_2^0 &:= Y_1^-(N_2, e^{-\mu_2}, 2^{-2\beta-8}) \\ \hat{N}_2^1 &:= Y_1^-(N_2, \mu_2 e^{-\mu_2}, 2^{-2\beta-8}) \\ \hat{N}_2^2 &:= Y_1^-(N_2, \omega_2 \mu_2^2 e^{-\mu_2}, 2^{-2\beta-8}) \end{aligned}$$

Then, when $\hat{\mathbf{M}}(\mathbf{M})$ satisfies Conditions 4 and 6, we define

$$\hat{\phi}_4(\mathbf{M}) := \hat{\phi}_3(\hat{\mathbf{M}}(\mathbf{M}), \vec{\mathbf{N}}), \quad (123)$$

otherwise, we define

$$\hat{\phi}_4(\mathbf{M}) := M. \quad (124)$$

Due to the definition, any element $\vec{\mathbf{N}} \in \Omega_1$ satisfies

$$\hat{\phi}_4(\mathbf{M}) \geq \hat{\phi}_2(\mathbf{M}, \vec{\mathbf{N}}), \quad (125)$$

i.e., the relation (42) holds. In summary, when the parameters $\mu_1, \mu_2, N, N_0, N_1$, and N_2 satisfy Condition 5, and when we choose the sacrifice bit length $S(\mathbf{M}) = \hat{\phi}_4(\mathbf{M}) + 2\beta + 5$ with the above choice of $\hat{\phi}_4(\mathbf{M})$, we obtain $\|\rho_{A,E} - \rho_{\text{ideal}}\|_1 \leq 2^{-\beta}$.

10. Improvement of evaluation

Up to the previous section, based on (45) given in Section 6, we give an evaluation of the universal composability with the finite-length setting. However, the above given evaluation can be improved by removing the square root for a part of probabilities in the following way. First, we change our definitions for estimates and the sacrifice bit-length in the following way.

$$\hat{J}^0(\bar{q}^0, N^0) := Y^-(N^0, \bar{q}^0, 2^{-\beta-6}) \quad (126)$$

$$\hat{J}^1(\bar{a}_\times^1, \bar{b}_\times^1, N^1) := Y^-(N^1, \bar{a}_\times^1 + \bar{b}_\times^1, 2^{-\beta-6}) \quad (127)$$

$$\hat{r}_\times^1(\bar{a}_\times^1, \bar{b}_\times^1, N^1) := p^+(\hat{J}^1(\bar{a}_\times^1 + \bar{b}_\times^1, N^1), \frac{\bar{b}_\times^1}{\bar{a}_\times^1 + \bar{b}_\times^1}, 2^{-2\beta-7}), \quad (128)$$

$$\hat{M}_0(M_0) := X^-(N_0, M_0, 2^{-\beta-6}), \quad (129)$$

$$\hat{M}_1(M_1) := X^-(N_1, M_1, 2^{-2\beta-7}), \quad (130)$$

$$\hat{M}_2(M_2) := X^+(N_2, M_2, 2^{-2\beta-7}), \quad (131)$$

$$\hat{M}_3(M_3) := X^+(N_1, M_3, 2^{-2\beta-7}), \quad (132)$$

$$\hat{N}_1^0 := Y_1^-(N_1, e^{-\mu_1}, 2^{-\beta-6}) \quad (133)$$

$$\hat{N}_1^1 := Y_1^-(N_1, \mu_1 e^{-\mu_1}, 2^{-\beta-6}) \quad (134)$$

$$\hat{N}_2^0 := Y_1^-(N_2, e^{-\mu_2}, 2^{-\beta-6}) \quad (135)$$

$$\hat{N}_2^1 := Y_1^-(N_2, \mu_2 e^{-\mu_2}, 2^{-\beta-6}) \quad (136)$$

$$\hat{N}_2^2 := Y_1^-(N_2, \omega_2 \mu_2^2 e^{-\mu_2}, 2^{-\beta-6}) \quad (137)$$

$$S(\mathbf{M}) := \hat{\phi}_4(\mathbf{M}) + 2\beta + 5. \quad (138)$$

When the signal intensity is μ_1 and the decoy intensity is μ_2 , we give the estimate $\hat{\mathbf{N}} = (\hat{N}^0, \hat{N}^1)$ as

$$\hat{N}^0 := Y_1^-(N, e^{-\mu_1}, 2^{-\beta-6}), \quad \hat{N}^1 := Y_1^-(N, \mu_1 e^{-\mu_1}, 2^{-\beta-6}).$$

Similarly, when the signal intensity is μ_2 and the decoy intensity is μ_1 , we give the estimate $\hat{\mathbf{N}} = (\hat{N}^0, \hat{N}^1)$ as

$$\hat{N}^0 := Y_1^-(N', e^{-\mu_2}, 2^{-\beta-6}), \quad \hat{N}^1 := Y_1^-(N', \mu_2 e^{-\mu_2}, 2^{-\beta-6}).$$

Further, Condition 5 is changed as

$$\begin{aligned} & Y_1^-(N_1, \mu_1 e^{-\mu_1}, 2^{-\beta-6}) Y_1^-(N_2, \omega_2 \mu_2^2 e^{-\mu_2}, 2^{-\beta-6}) \\ & > (N_1 - Y_1^-(N_1, e^{-\mu_1}, 2^{-\beta-6}) - Y_1^-(N_1, \mu_1 e^{-\mu_1}, 2^{-\beta-6})) Y_1^+(N_2, \mu_2 e^{-\mu_2}, 2^{-\beta-6}), \\ & Y_1^-(N_2, \omega_2 \mu_2^2 e^{-\mu_2}, 2^{-\beta-6}) Y_1^-(N_1, e^{-\mu_1}, 2^{-\beta-6}) \\ & > (N_1 - Y_1^-(N_1, e^{-\mu_1}, 2^{-\beta-6}) - Y_1^-(N_1, \mu_1 e^{-\mu_1}, 2^{-\beta-6})) Y_1^+(N_2, e^{-\mu_2}, 2^{-\beta-6}), \\ & \frac{Y_1^-(N_2, \omega_2 \mu_2^2 e^{-\mu_2}, 2^{-\beta-6})}{N_1 - Y_1^-(N_1, e^{-\mu_1}, 2^{-\beta-6}) - Y_1^-(N_1, \mu_1 e^{-\mu_1}, 2^{-\beta-6})} + \frac{Y_1^-(N_2, e^{-\mu_2}, 2^{-\beta-6})}{Y_1^+(N_1, e^{-\mu_1}, 2^{-\beta-6})} \\ & > \frac{2Y_1^+(N_2, \mu_2 e^{-\mu_2}, 2^{-\beta-6})}{Y_1^-(N_1, \mu_1 e^{-\mu_1}, 2^{-\beta-6})}. \end{aligned}$$

Then, we obtain the following theorem

Theorem 5 *Under the above definition, when Condition 5 holds, we obtain*

$$\|\rho_{A,E} - \rho_{\text{ideal}}\|_1 \leq 2^{-\beta}. \quad (139)$$

Now, we will show the above theorem. For this purpose, we treat Formula (2) more deeply. Assume that $\mathbf{s} = (s_1, \dots, s_{N+N'+N_0+N_1+N_2})$ is the sequence of the indicator of the kinds of the generated state among all of $N + N' + N_0 + N_1 + N_2$ sent states. For example, if the i -th received state is the vacuum state, s_i is 0. if the i -th received state is the single-photon state, s_i is 1. if the i -th received state is the state ρ_2 , s_i is 2. Otherwise, s_i is 3. We apply (2) to the case when \mathbf{s} and \mathbf{N} are fixed, we obtain

$$\|\rho_{A,E|\mathbf{s},\mathbf{N}} - \rho_{\text{ideal}}|_{\mathbf{s},\mathbf{N}}\|_1 \leq 2\sqrt{2}\sqrt{P_{ph|\mathbf{s},\mathbf{N}}}, \quad (140)$$

where $\rho_{A,E|\mathbf{s},\mathbf{N}}$, $\rho_{\text{ideal}}|_{\mathbf{s},\mathbf{N}}$ and $P_{ph|\mathbf{s},\mathbf{N}}$ are the final true composite state, the ideal final state, and the averaged virtual phase error probability with conditioned \mathbf{s}, \mathbf{N} . Hence,

the final true composite state $\rho_{A,E}$ and the ideal final state ρ_{ideal} are written as

$$\begin{aligned}\rho_{A,E} &= \sum_{\mathbf{s}, \mathbf{N}} P(\mathbf{s}, \mathbf{N}) |\mathbf{s}, \mathbf{N}\rangle \langle \mathbf{s}, \mathbf{N}| \otimes \rho_{A,E|\mathbf{s}, \mathbf{N}}, \\ \rho_{\text{ideal}} &= \sum_{\mathbf{s}, \mathbf{N}} P(\mathbf{s}, \mathbf{N}) |\mathbf{s}, \mathbf{N}\rangle \langle \mathbf{s}, \mathbf{N}| \otimes \rho_{\text{ideal}|\mathbf{s}, \mathbf{N}}.\end{aligned}$$

Hence, when Ω is a set concerning \mathbf{s}, \mathbf{N} ,

$$\begin{aligned}\|\rho_{A,E} - \rho_{\text{ideal}}\|_1 &= \sum_{\mathbf{s}, \mathbf{N}} P(\mathbf{s}, \mathbf{N}) \|\rho_{A,E|\mathbf{s}, \mathbf{N}} - \rho_{\text{ideal}|\mathbf{s}, \mathbf{N}}\|_1 \\ &\leq \sum_{\mathbf{s}, \mathbf{N}} P(\mathbf{s}, \mathbf{N}) \min\{2\sqrt{2}\sqrt{P_{ph|\mathbf{s}, \mathbf{N}}}, 2\} \\ &\leq 2P(\Omega^c) + 2\sqrt{2} \sqrt{\sum_{\mathbf{s}, \mathbf{N} \in \Omega} P(\mathbf{s}, \mathbf{N}) P_{ph|\mathbf{s}, \mathbf{N}}}.\end{aligned}\tag{141}$$

When the signal intensity is μ_1 and the decoy intensity is μ_2 , for the analysis, we replace the definition of Ω_1 as the following way.

$$N^0 \in [Y_1^-(N, e^{-\mu_1}, 2^{-\beta-6}), Y_1^+(N, e^{-\mu_1}, 2^{-\beta-6})]\tag{142}$$

$$N^1 \in [Y_1^-(N, \mu_1 e^{-\mu_1}, 2^{-\beta-6}), Y_1^+(N, \mu_1 e^{-\mu_1}, 2^{-\beta-6})]\tag{143}$$

$$N_1^0 \in [Y_1^-(N_1, e^{-\mu_1}, 2^{-\beta-6}), Y_1^+(N_1, e^{-\mu_1}, 2^{-\beta-6})]\tag{144}$$

$$N_1^1 \in [Y_1^-(N_1, \mu_1 e^{-\mu_1}, 2^{-\beta-6}), Y_1^+(N_1, \mu_1 e^{-\mu_1}, 2^{-\beta-6})]\tag{145}$$

$$N_2^0 \in [Y_1^-(N_2, e^{-\mu_2}, 2^{-\beta-6}), Y_1^+(N_2, e^{-\mu_2}, 2^{-\beta-6})]\tag{146}$$

$$N_2^1 \in [Y_1^-(N_2, \mu_2 e^{-\mu_2}, 2^{-\beta-6}), Y_1^+(N_2, \mu_2 e^{-\mu_2}, 2^{-\beta-6})]\tag{147}$$

$$N_2^2 \in [Y_1^-(N_2, \omega_2 \mu_2^2 e^{-\mu_2}, 2^{-\beta-6}), Y_1^+(N_2, \omega_2 \mu_2^2 e^{-\mu_2}, 2^{-\beta-6})]\tag{148}$$

When the signal intensity is μ_2 and the decoy intensity is μ_1 , (142) and (143) are replaced by

$$N^0 \in [Y_1^-(N', e^{-\mu_1}, 2^{-\beta-6}), Y_1^+(N', e^{-\mu_1}, 2^{-\beta-6})]\tag{149}$$

$$N^1 \in [Y_1^-(N', \mu_1 e^{-\mu_1}, 2^{-\beta-6}), Y_1^+(N', \mu_1 e^{-\mu_1}, 2^{-\beta-6})].\tag{150}$$

We choose the set Ω as

$$\Omega := \Omega_1 \cap \{J^0 \geq Y^-(N^0, \bar{q}^0, 2^{-\beta-6})\} \cap \{J^1 \geq Y^-(N^1, \bar{a}_x^1 + \bar{b}_x^1, 2^{-\beta-6})\} \cap \{\bar{M}_0 \geq \hat{M}_0(M_0)\}.$$

Hence, using (141), we obtain

$$\|\rho_{A,E} - \rho_{\text{ideal}}\|_1 \leq 2 \cdot (7 \cdot 2 + 3) \cdot 2^{-\beta-6} + 2\sqrt{2} \sqrt{\sum_{\mathbf{s}, \mathbf{N} \in \Omega} P(\mathbf{s}, \mathbf{N}) P_{ph|\mathbf{s}, \mathbf{N}}}.\tag{151}$$

Since $\hat{\phi}_4(\mathbf{M}) \geq \hat{\phi}_2(\hat{\mathbf{M}}(\mathbf{M}), \vec{\mathbf{N}})$ for $\vec{\mathbf{N}} \in \Omega_1$, as is shown in (125), due to (95), (96), and (57), we obtain

$$\begin{aligned}\Omega \cap \{\hat{\phi}_4(\mathbf{M}) < \phi(\mathbf{J})\} &\subset \Omega \cap \{\hat{\phi}_2(\mathbf{M}, \vec{\mathbf{N}}) < \phi(\mathbf{J})\} \\ &\subset \Omega \cap \left(\{\phi(\mathbf{J}) > \hat{\phi}_1(\mathbf{N}, \bar{q}^0, \bar{a}_x^1, \bar{b}_x^1)\} \cup \{\hat{\phi}_3(\bar{\mathbf{M}}, \vec{\mathbf{N}}) > \hat{\phi}_2(\mathbf{M}, \vec{\mathbf{N}})\} \right) \\ &\subset \Omega \cap \left(\{J^0 \leq Y^-(N^0, \bar{q}^0, 2^{-\beta-6})\} \cup \{J^1 \leq Y^-(N^1, \bar{a}_x^1, \bar{b}_x^1, 2^{-\beta-6})\} \right)\end{aligned}$$

$$\begin{aligned}
& \cup \left\{ \frac{J^3}{J^1} \geq p^+(J^1, \frac{\bar{b}_x^1}{\bar{a}_x^1 + \bar{b}_x^1}, 2^{-2\beta-7}) \right\} \\
& \cup \{ \bar{M}_0 < \hat{M}_0(M_0) \} \cup \{ \bar{M}_1 < \hat{M}_1(M_1) \} \cup \{ \bar{M}_2 > \hat{M}_2(M_2) \} \cup \{ \bar{M}_3 > \hat{M}_3(M_3) \} \\
& \subset \Omega \cap \left(\left\{ \frac{J^3}{J^1} \geq p^+(J^1, \frac{\bar{b}_x^1}{\bar{a}_x^1 + \bar{b}_x^1}, 2^{-2\beta-7}) \right\} \right. \\
& \quad \left. \cup \{ \bar{M}_1 < \hat{M}_1(M_1) \} \cup \{ \bar{M}_2 > \hat{M}_2(M_2) \} \cup \{ \bar{M}_3 > \hat{M}_3(M_3) \} \right) \\
& \subset \left\{ \frac{J^3}{J^1} \geq p^+(J^1, \frac{\bar{b}_x^1}{\bar{a}_x^1 + \bar{b}_x^1}, 2^{-2\beta-7}) \right\} \cup \{ \bar{M}_1 < \hat{M}_1(M_1) \} \cup \{ \bar{M}_2 > \hat{M}_2(M_2) \} \cup \{ \bar{M}_3 > \hat{M}_3(M_3) \}
\end{aligned}$$

Hence, due to (138),

$$\begin{aligned}
& \sum_{\mathbf{s}, \mathbf{N} \in \Omega} P(\mathbf{s}, \mathbf{N}) P_{ph|\mathbf{s}, \mathbf{N}} \leq \Pr \Omega \cap \{ \hat{\phi}_4(\mathbf{M}) < \phi(\mathbf{J}) \} + 2^{-2\beta-5} \\
& \leq 4 \cdot 2^{-2\beta-7} + 2^{-2\beta-5} = 2^{-2\beta-4}.
\end{aligned}$$

Thus, using (151), we obtain

$$\| \rho_{A,E} - \rho_{\text{ideal}} \|_1 \leq 2 \cdot 17 \cdot 2^{-\beta-6} + 2\sqrt{2} \cdot 2^{\frac{-2\beta-4}{2}} \leq 2^{-\beta} \quad (153)$$

because $2\sqrt{2} + \frac{17}{16} (\cong 3.89) \leq 4$. Indeed, in order to put out a probability from the square root, the event corresponding to the probability must be defined by \mathbf{s}, \mathbf{N} . Hence, the probabilities corresponding to the sets $\{ \frac{J^3}{J^1} \geq p^+(J^1, \frac{\bar{b}_x^1}{\bar{a}_x^1 + \bar{b}_x^1}, 2^{-2\beta-7}) \}$, $\{ \bar{M}_1 < \hat{M}_1(M_1) \}$, $\{ \bar{M}_2 > \hat{M}_2(M_2) \}$, and $\{ \bar{M}_3 > \hat{M}_3(M_3) \}$ cannot be put out from the square root.

In summary, when the parameters μ_1 , μ_2 , N , N_0 , N_1 , and N_2 satisfy Condition 5 modified in this section, and when we choose the sacrifice bit length $S(\mathbf{M}) = \hat{\phi}_4(\mathbf{M}) + 2\beta + 5$ by using the choice of $\hat{\phi}_4(\mathbf{M})$ given in (123) and (124) and the parameters given in this section, we obtain $\| \rho_{A,E} - \rho_{\text{ideal}} \|_1 \leq 2^{-\beta}$.

11. Asymptotic optimization of decoy intensity μ_1

11.1. Asymptotic expansion of $\hat{b}_x^1(\hat{\mathbf{M}}(\bar{\mathbf{M}}), \hat{\mathbf{N}})$

Next, for a given signal intensity, we consider the minimization of the sacrifice bit length concerning the decoy intensity by taking into account statistical fluctuations with an asymptotic setting. As is already discussed in Section 8, it is better to make the decoy intensity smaller than the signal intensity. Hence, in the following, we consider the case when the signal intensity is μ_2 and the decoy intensity is μ_1 with $\mu_1 < \mu_2$. While the optimal decoy intensity μ_1 is zero in Theorem 3, this argument does not take into account statistical fluctuations. Hence, we need to optimize the decoy intensity μ_1 with the choice of $\hat{\mathbf{N}}$ and $\hat{\mathbf{M}}(\bar{\mathbf{M}})$.

Due to Theorem 3, the optimal decoy intensity converges to zero when N' , N_0 , N_1 , and N_2 go to infinity. Now, based on the model (104), we consider the case when $N_0 = c_0 N'$, $N_i = c_i N'$, $M_0 = p_0 N_0$, $M_i = ((1-s)(1-e^{-\alpha\mu_i}) + p_0/2)N_i$, and

$M_{i+2} = (s(1 - e^{-\alpha\mu_i}) + p_0/2)N_i$ for $i = 1, 2$. Choosing $\mu_1 = \frac{\lambda}{N'^{1/4}}$, we derive the asymptotic expansion of $\hat{\phi}_4(\bar{\mathbf{M}})/N'$ up to the order $\frac{1}{N'^{1/4}}$. The purpose of this section is to optimize λ .

For this optimization, we have to treat two kinds of factors. One is the speed of the convergence to the asymptotic rate. The other is the asymptotic rate. With respect to the finite-length key generation rate, both factors add negative terms to the asymptotic key generation rate when the parameters \bar{a}_\times^1 and \bar{b}_\times^1 are perfectly estimated. These terms appear as terms with smaller orders than N . The absolute value of the term by the first factor is monotonely increasing with respect to the decoy intensity, and that by the second factor is monotonely decreasing with respect to the decoy intensity. Hence, we have to address the trade-off between two terms.

Indeed, if we choose a smaller order than $\frac{\lambda}{N'^{1/4}}$ for the decoy intensity μ_1 , the absolute value of the term by the first factor has a larger order than that by the second factor. That is, the optimal coefficient is infinity. If we choose a larger order than $\frac{\lambda}{N'^{1/4}}$ for the decoy intensity μ_1 , the absolute value of the term by the second factor has a larger order than that by the first factor. That is, the optimal coefficient is zero. In both cases, the asymptotic expansion is not valid with such extremal cases. In order to address the above trade-off, we need to choose the order of the decoy intensity μ_1 such that the both terms has the same order. This requirement is satisfied when the decoy intensity behaves as $\frac{\lambda}{N'^{1/4}}$.

In the following, we first treat the asymptotic expansion of $\hat{b}_\times^1(\hat{\mathbf{M}}(\bar{\mathbf{M}}), \hat{\mathbf{N}})$ up to the order $\frac{1}{N'^{1/4}}$. Using the discussion in Appendix C, we obtain the asymptotic expansions of $\hat{M}_0(\bar{M}_0)/N'$, $\hat{M}_1(\bar{M}_1)/N'$, $\hat{M}_2(\bar{M}_2)/N'$, $\hat{M}_3(\bar{M}_3)/N'$, \hat{N}_1^0/N' , \hat{N}_1^1/N' , \hat{N}_2^0/N' , \hat{N}_2^1/N' , \hat{N}_2^2/N' , \hat{N}^0/N' , and \hat{N}^1/N' as follows.

$$\begin{aligned}\hat{M}_0(\bar{M}_0)/N' &= c_0 p_0 + \sqrt{c_0 p_0 (1 - p_0)} x_0 \frac{1}{N'^{1/2}} + o\left(\frac{1}{N'^{1/2}}\right), \\ \hat{M}_1(\bar{M}_1)/N' &= c_1 \left(\frac{p_0}{2} + (1 - s)(1 - e^{-\alpha\mu_1})\right) - \sqrt{c_1 \frac{p_0}{2} \left(1 - \frac{p_0}{2}\right)} x_1 \frac{1}{N'^{1/2}} + o\left(\frac{1}{N'^{1/2}}\right), \\ \hat{M}_2(\bar{M}_2)/N' &= c_2 \left(\frac{p_0}{2} + (1 - s)(1 - e^{-\alpha\mu_2})\right) + o\left(\frac{1}{N'^{1/4}}\right), \\ \hat{M}_3(\bar{M}_3)/N' &= c_1 \left(\frac{p_0}{2} + s(1 - e^{-\alpha\mu_1})\right) + \sqrt{c_1 \frac{p_0}{2} \left(1 - \frac{p_0}{2}\right)} x_1 \frac{1}{N'^{1/2}} + o\left(\frac{1}{N'^{1/2}}\right), \\ \hat{N}_1^0/N' &= c_1 e^{-\mu_1} + o\left(\frac{1}{N'^{1/2}}\right) \\ \hat{N}_1^1/N' &= c_1 \mu_1 e^{-\mu_1} + o\left(\frac{1}{N'^{1/2}}\right) = c_1 \mu_1 e^{-\mu_1} \left(1 + o\left(\frac{1}{N'^{1/2}}\right)\right) \\ \hat{N}_1^2/N' &= c_1 \omega_2 \mu_1^2 e^{-\mu_1} + o\left(\frac{1}{N'^{1/2}}\right) \\ \hat{N}_2^0/N' &= c_2 e^{-\mu_2} + o\left(\frac{1}{N'^{1/4}}\right) \\ \hat{N}_2^1/N' &= c_2 \mu_2 e^{-\mu_2} + o\left(\frac{1}{N'^{1/4}}\right)\end{aligned}$$

$$\begin{aligned}\hat{N}_2^2/N' &= c_2\omega_2\mu_2^2e^{-\mu_1} + o(\frac{1}{N'^{1/4}}) \\ \hat{N}^0/N' &= e^{-\mu_2} + o(\frac{1}{N'^{1/4}})\end{aligned}\tag{154}$$

$$\hat{N}^1/N' = \mu_2e^{-\mu_2} + o(\frac{1}{N'^{1/4}}),\tag{155}$$

where x_1 and x_0 are defined as $2^{-\beta-6} = \int_{x_0}^{\infty} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} dx$ and $2^{-2\beta-7} = \int_{x_1}^{\infty} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} dx$ ¶. Relation to the latter discussion, we need asymptotic expansions for $\bar{M}_0/N', \bar{M}_1/N', \bar{M}_3/N', N_1^0/N', N_1^1/N', N_1^2/N'$ up to the order $\frac{1}{N'^{1/2}}$.

First, we treat the asymptotic expansion for $\hat{b}_x^1(\bar{\mathbf{M}}, \bar{\mathbf{N}})$

$$\begin{aligned}& \frac{\bar{M}_3 - \bar{M}_0N_1^0/2N_0}{N'} \\ &= (c_1(\frac{p_0}{2} + s(1 - e^{-\alpha\mu_1})) + \sqrt{c_1\frac{p_0}{2}(1 - \frac{p_0}{2})}x_0\frac{1}{N'^{1/2}} + o(\frac{1}{N'^{1/2}})) \\ & \quad - \frac{1}{2c_0}(c_0p_0 + \sqrt{c_0p_0(1 - p_0)}x_1\frac{1}{N'^{1/2}} + o(\frac{1}{N'^{1/2}}))(c_1e^{-\mu_1} + o(\frac{1}{N'^{1/2}})) \\ &= c_1(\frac{p_0}{2} + s(1 - e^{-\alpha\mu_1})) + \sqrt{c_1\frac{p_0}{2}(1 - \frac{p_0}{2})}x_1\frac{1}{N'^{1/2}} \\ & \quad - c_1e^{-\mu_1}\frac{p_0}{2} - \frac{c_1}{2}e^{-\mu_1}\sqrt{p_0(1 - p_0)/c_0x_0}\frac{1}{N'^{1/2}} + o(\frac{1}{N'^{1/2}}) \\ &= c_1(\frac{p_0}{2}(1 - e^{-\mu_1}) + s(1 - e^{-\alpha\mu_1})) + \sqrt{c_1\frac{p_0}{2}(1 - \frac{p_0}{2})}x_1\frac{1}{N'^{1/2}} \\ & \quad - \frac{c_1}{2}e^{-\mu_1}\sqrt{p_0(1 - p_0)/c_0x_0}\frac{1}{N'^{1/2}} + o(\frac{1}{N'^{1/2}}) \\ &= c_1(\frac{p_0}{2}(1 - e^{-\mu_1}) + s(1 - e^{-\alpha\mu_1})) \\ & \quad + (\sqrt{c_1\frac{p_0}{2}(1 - \frac{p_0}{2})}x_1 - \frac{c_1}{2}e^{-\mu_1}\sqrt{p_0(1 - p_0)/c_0x_0})\frac{1}{N'^{1/2}} + o(\frac{1}{N'^{1/2}}).\end{aligned}\tag{156}$$

Since the order of $c_1\mu_1e^{-\mu_1}$ is $\frac{1}{N'^{1/4}}$, using (114), we obtain

$$\begin{aligned}& \frac{\bar{M}_3 - \bar{M}_0N_1^0/2N_0}{c_1\mu_1e^{-\mu_1}N'} \\ &= \frac{\frac{p_0}{2}(e^{\mu_1} - 1) + s(e^{\mu_1} - e^{(1-\alpha)\mu_1})}{\mu_1} \\ & \quad + \frac{\sqrt{c_1\frac{p_0}{2}(1 - \frac{p_0}{2})}x_1 - \frac{c_1}{2}e^{-\mu_1}\sqrt{p_0(1 - p_0)/c_0x_0}}{c_1\mu_1e^{-\mu_1}}\frac{1}{N'^{1/2}} + o(\frac{1}{N'^{1/4}}) \\ &= s\alpha + \frac{p_0}{2} + (F_1\lambda + \frac{G_1x_1 - G_2x_0}{\lambda})\frac{1}{N'^{1/4}} + o(\frac{1}{N'^{1/4}}),\end{aligned}$$

¶ If we do not employ the central limit theorem, we can employ Chernoff bound, i.e., \tilde{p}^\pm given in Appendix A. In this case, as is discussed in Appendix C, x_1 and x_0 are defined to be $\sqrt{\frac{2(2\beta+7)}{\log e}}$ and $\sqrt{\frac{2(\beta+6)}{\log e}}$.

where

$$F_1 := \frac{1}{2}s(1 - (1 - \alpha)^2) + \frac{p_0}{2}, \quad G_1 := \sqrt{\frac{p_0}{2c_1}(1 - \frac{p_0}{2})}, \quad G_2 := \frac{1}{2}\sqrt{p_0(1 - p_0)/c_0}.$$

Hence,

$$\begin{aligned} \hat{b}_\times^1(\bar{\mathbf{M}}, \vec{\mathbf{N}}) &= \frac{\bar{M}_3 - \bar{M}_0 N_1^0 / 2N_0}{N_1^1} \\ &= s\alpha + \frac{p_0}{2} + (F_1\lambda + \frac{G_1 x_1 - G_2 x_0}{\lambda}) \frac{1}{N'^{1/4}} + o(\frac{1}{N'^{1/4}}). \end{aligned}$$

11.2. Asymptotic expansion of $\hat{b}_\times^1(\bar{\mathbf{M}}, \vec{\mathbf{N}})$

Next, we treat the asymptotic expansion of $\hat{a}_\times^1(\bar{\mathbf{M}}, \vec{\mathbf{N}})$ up to the order $\frac{1}{N'^{1/4}}$. Since $N_1^1/N' = O(\frac{1}{N'^{1/4}})$, and $N_1^2/N' = O(\frac{1}{N'^{1/2}})$,

$$\frac{N_1^1 N_2^2 - N_2^1 N_1^2}{N'^2} = c_1 c_2 \omega_2 e^{-\mu_1 - \mu_2} \mu_1 \mu_2 (\mu_2 - \mu_1) + o(\frac{1}{N'^{1/2}})$$

Since $c_1 c_2 \omega_2 e^{-\mu_1 - \mu_2} \mu_1 \mu_2 (\mu_2 - \mu_1) = O(\frac{1}{N'^{1/4}})$,

$$N_1^1 N_2^2 - N_2^1 N_1^2 = c_1 c_2 \omega_2 e^{-\mu_1 - \mu_2} \mu_1 \mu_2 (\mu_2 - \mu_1) N'^2 (1 + o(\frac{1}{N'^{1/4}})). \quad (157)$$

Similar to (156), we have

$$\begin{aligned} &\frac{\bar{M}_1 - \bar{M}_0 N_1^0 / 2N_0}{N'} \\ &= c_1 \left(\frac{p_0}{2} (1 - e^{-\mu_1}) + (1 - s)(1 - e^{-\alpha\mu_1}) \right) \\ &\quad - \left(\sqrt{c_1 \frac{p_0}{2} (1 - \frac{p_0}{2})} x_1 + \frac{c_1}{2} e^{-\mu_1} \sqrt{p_0(1 - p_0)/c_0 x_0} \right) \frac{1}{N'^{1/2}} + o(\frac{1}{N'^{1/2}}), \end{aligned} \quad (158)$$

whose order is $O(\frac{1}{N'^{1/4}})$ because $c_1(p_0(1 - e^{-\mu_1}) + 1 - e^{-\alpha\mu_1}) = O(\frac{1}{N'^{1/4}})$. Since $N_1^2 = O(\frac{1}{N'^{1/4}})$, we obtain

$$\begin{aligned} &\frac{N_2^2(\bar{M}_1 - \bar{M}_0 N_1^0 / 2N_0) - N_1^2(\bar{M}_2 - \bar{M}_0 N_2^0 / 2N_0)}{N'^2} \\ &= c_2 \omega_2 \mu_2^2 e^{-\mu_2} c_1 \left(\frac{p_0}{2} (1 - e^{-\mu_1}) + (1 - s)(1 - e^{-\alpha\mu_1}) \right) \\ &\quad - c_2 \omega_2 \mu_2^2 e^{-\mu_2} \left(\sqrt{c_1 \frac{p_0}{2} (1 - \frac{p_0}{2})} x_1 + \frac{c_1}{2} e^{-\mu_1} \sqrt{p_0(1 - p_0)/c_0 x_0} \right) \frac{1}{N'^{1/2}} \\ &\quad - c_1 \omega_2 \mu_1^2 e^{-\mu_1} c_2 \left(\frac{p_0}{2} (1 - e^{-\mu_2}) + (1 - s)(1 - e^{-\alpha\mu_2}) \right) + o(\frac{1}{N'^{1/2}}). \end{aligned} \quad (159)$$

On the other hand,

$$\begin{aligned} &\frac{c_2 \omega_2 \mu_2^2 e^{-\mu_2} \left(\sqrt{c_1 \frac{p_0}{2} (1 - \frac{p_0}{2})} x_1 + \frac{c_1}{2} e^{-\mu_1} \sqrt{p_0(1 - p_0)/c_0 x_0} \right)}{c_1 c_2 \omega_2 e^{-\mu_1 - \mu_2} \mu_1 \mu_2 (\mu_2 - \mu_1)} \frac{1}{N'^{1/2}} \\ &= \frac{G_1 x_1 + G_2 x_0}{\lambda} \frac{1}{N'^{1/4}} + o(\frac{1}{N'^{1/4}}). \end{aligned}$$

Using the quantities

$$F_2(\mu_2) := (1 - s + \frac{p_0}{2}) \frac{e^{\mu_2} - 1 - \mu_2 - \mu_2^2/2}{\mu_2^2} - (1 - s) \frac{e^{(1-\alpha)\mu_2} - 1 - (1-\alpha)\mu_2 - (1-\alpha)^2\mu_2^2/2}{\mu_2^2},$$

we have

$$\begin{aligned} & \frac{\mu_2^2 e^{-\mu_2} (\frac{p_0}{2}(1 - e^{-\mu_1}) + (1-s)(1 - e^{-\alpha\mu_1})) - \mu_1^2 e^{-\mu_1} (\frac{p_0}{2}(1 - e^{-\mu_2}) + (1-s)(1 - e^{-\alpha\mu_2}))}{e^{-\mu_1-\mu_2} \mu_1 \mu_2 (\mu_2 - \mu_1)} \\ &= \frac{p_0}{2} + (1-s)\alpha - \sum_{n=3}^{\infty} \frac{(1-s)(1 - (1-\alpha)^n) + \frac{p_0}{2}}{n!} \left(\sum_{m=0}^{n-3} \mu_1^{m+1} \mu_2^{n-2-m} \right) \\ &= \frac{p_0}{2} + (1-s)\alpha - \sum_{n=3}^{\infty} \frac{(1-s)(1 - (1-\alpha)^n) + \frac{p_0}{2}}{n!} \mu_1 \mu_2^{n-2} + o(\frac{1}{N'^{1/4}}) \\ &= \frac{p_0}{2} + (1-s)\alpha - \mu_1 F_2(\mu_2) + o(\frac{1}{N'^{1/4}}) \\ &= \frac{p_0}{2} + (1-s)\alpha - \frac{\lambda}{N'^{1/4}} F_2(\mu_2) + o(\frac{1}{N'^{1/4}}). \end{aligned} \quad (160)$$

Due to the relation $c_1 c_2 \omega_2 e^{-\mu_1-\mu_2} \mu_1 \mu_2 (\mu_2 - \mu_1) = O(\frac{1}{N'^{1/4}})$, we obtain

$$\begin{aligned} & \frac{N_2^2 (\bar{M}_1 - \bar{M}_0 N_1^0 / N_0) - N_1^2 (\bar{M}_2 - \bar{M}_0 N_2^0 / N_0)}{c_1 c_2 \omega_2 e^{-\mu_1-\mu_2} \mu_1 \mu_2 (\mu_2 - \mu_1) N'^2} \\ &= \frac{p_0}{2} + (1-s)\alpha - (F_2(\mu_2)\lambda + \frac{G_1 x_1 + G_2 x_0}{\lambda}) \frac{1}{N'^{1/4}} + o(\frac{1}{N'^{1/4}}). \end{aligned} \quad (161)$$

Therefore, combining (157) and (161), we obtain

$$\begin{aligned} \hat{a}_\times^1(\bar{\mathbf{M}}, \bar{\mathbf{N}}) &= \frac{N_2^2 (\bar{M}_1 - \bar{M}_0 N_1^0 / N_0) - N_1^2 (\bar{M}_2 - \bar{M}_0 N_2^0 / N_0)}{N_1^1 N_2^2 - N_2^1 N_1^2} \\ &= \frac{p_0}{2} + (1-s)\alpha - [\lambda F_2(\mu_2) + \frac{G_1 x_1 + G_2 x_0}{\lambda}] \frac{1}{N'^{1/4}} + o(\frac{1}{N'^{1/4}}). \end{aligned}$$

11.3. Asymptotic minimization of sacrifice bit length

When a signal pulse μ_2 is fixed, for minimization of sacrifice bit length, it is sufficient to maximize $(\hat{a}_\times^1(\bar{\mathbf{M}}, \bar{\mathbf{N}}) + \hat{b}_\times^1(\bar{\mathbf{M}}, \bar{\mathbf{N}}))(1 - h(\frac{\hat{b}_\times^1(\bar{\mathbf{M}}, \bar{\mathbf{N}})}{\hat{a}_\times^1(\bar{\mathbf{M}}, \bar{\mathbf{N}}) + \hat{b}_\times^1(\bar{\mathbf{M}}, \bar{\mathbf{N}})}))$ because $N^{0'}$ and $N^{1'}$ do not depend of the decoy intensity μ_1 . Using the partial derivatives of the function $(x, y) \mapsto (x + y)(1 - h(\frac{y}{x+y}))$, we obtain

$$\begin{aligned} & (\hat{a}_\times^1(\bar{\mathbf{M}}, \bar{\mathbf{N}}) + \hat{b}_\times^1(\bar{\mathbf{M}}, \bar{\mathbf{N}}))(1 - h(\frac{\hat{b}_\times^1(\bar{\mathbf{M}}, \bar{\mathbf{N}})}{\hat{a}_\times^1(\bar{\mathbf{M}}, \bar{\mathbf{N}}) + \hat{b}_\times^1(\bar{\mathbf{M}}, \bar{\mathbf{N}})})) \\ &= (p_0 + \alpha)(1 - h(\frac{p_0/2 + s\alpha}{p_0 + \alpha})) + (1 + \log \frac{p_0/2 + s\alpha}{p_0 + \alpha}) [\lambda F_1 + \frac{G_1 x_1 - G_2 x_0}{\lambda}] \frac{1}{N'^{1/4}} \\ & \quad - (1 + \log \frac{p_0/2 + (1-s)\alpha}{p_0 + \alpha}) [\lambda F_2(\mu_2) + \frac{G_1 x_1 + G_2 x_0}{\lambda}] \frac{1}{N'^{1/4}} + o(\frac{1}{N'^{1/4}}) \\ &= (p_0 + \alpha)(1 - h(\frac{p_0/2 + s\alpha}{p_0 + \alpha})) - (C_1 \lambda + \frac{C_2}{\lambda}) \frac{1}{N'^{1/4}} + o(\frac{1}{N'^{1/4}}), \end{aligned}$$

where

$$C_1 := (1 + \log \frac{p_0/2 + (1-s)\alpha}{p_0 + \alpha})F_2(\mu_2) - (1 + \log \frac{p_0/2 + s\alpha}{p_0 + \alpha})F_1$$

$$C_2 := (1 + \log \frac{p_0/2 + (1-s)\alpha}{p_0 + \alpha})(G_1x_1 + G_2x_0) - (1 + \log \frac{p_0/2 + s\alpha}{p_0 + \alpha})(G_1x_1 - G_2x_0).$$

Therefore, it follows from (154), (155), and the above argument that the sacrifice bit length is

$$M - N'e^{-\mu_2} - N'\mu_2e^{-\mu_2}(p_0 + \alpha)(1 - h(\frac{p_0/2 + s\alpha}{p_0 + \alpha})) + \mu_2e^{-\mu_2}(C_1\lambda + \frac{C_2}{\lambda})N'^{3/4} + o(N'^{3/4})$$

$$= M - N'e^{-\mu_2} - N'\mu_2e^{-\mu_2}(p_0 + \alpha)(1 - h(\frac{p_0/2 + s\alpha}{p_0 + \alpha})) + \mu_2e^{-\mu_2}(C_1\mu_1 + \frac{C_2}{\mu_1\sqrt{N'}})N' + o(N'^{3/4}).$$
(162)

Due to the relation between arithmetic mean and geometric mean, the maximum of the coefficient of the order $\frac{1}{N'^{1/4}}$, can be attained with $\lambda_* := \sqrt{\frac{C_2}{C_1}}$, i.e., $\mu_1 = \sqrt{\frac{C_2}{C_1}}N'^{-1/4}$. Therefore, the minimum sacrifice bit length with a fixed signal pulse μ_2 is

$$M - N'e^{-\mu_2} - N'\mu_2e^{-\mu_2}(p_0 + \alpha)(1 - h(\frac{p_0/2 + s\alpha}{p_0 + \alpha})) + 2\mu_2e^{-\mu_2}\sqrt{C_1C_2}N'^{3/4} + o(N'^{3/4}).$$

12. Analysis when the intensities are not fixed

12.1. Case when the intensities μ_1 and μ_2 obey the independent and identical distribution

Unfortunately, many realized quantum key distribution system have fluctuation for the intensities. In Section 9, we have derived the secure sacrifice bit length when the breakdowns of N_1 pulses and N_2 pulses obey the Poisson distribution. However, when the intensities have fluctuation, we have to derive the sacrifice bit length by taking into account this factor. That is, we need to discuss the distribution for \vec{N} in the way different from that in Section 9. In Subsection 8.3, we have already discussed the case concerning the asymptotic key generation rate. In this section, we discuss the sacrifice bit length in the same setting.

Indeed, the definition of $\hat{\phi}_2(\mathbf{M}, \vec{N})$ does not depend on the distribution for \vec{N} . Hence, since it is not needed to change the definition of $\hat{\phi}_2(\mathbf{M}, \vec{N})$, the relation (34) holds without any modification. Thus, we need to modify the definitions of $\hat{\phi}_4(\mathbf{M})$ and the set Ω_1 giving the fluctuation of \vec{N} so that the relation (42) holds. Further, since the definition of ρ_2 given in (10) depends on the intensity μ_1 , we need to modify the definition of ρ_2 properly.

In the following, we assume that the intensities μ_1 and μ_2 independently obey independent and identical distributions satisfying the following condition. For any integer $n \geq 3$, the relation

$$\mathbb{E}[e^{-\mu_2}\mu_2^n]\mathbb{E}[e^{-\mu_1}\mu_1^2] \geq \mathbb{E}[e^{-\mu_1}\mu_1^n]\mathbb{E}[e^{-\mu_2}\mu_2^2]$$

holds, where E denotes the expectation. Under the above assumption, we have expansions for two kinds of pulses.

$$\sum_{n=0}^{\infty} \frac{E[e^{-\mu_1} \mu_1^n]}{n!} |n\rangle \langle n| = E[e^{-\mu_1}] |0\rangle \langle 0| + E[e^{-\mu_1} \mu_1] |1\rangle \langle 1| + E[e^{-\mu_1} \mu_1^2] \omega_2 \rho_2 \quad (163)$$

$$\sum_{n=0}^{\infty} \frac{E[e^{-\mu_2} \mu_2^n]}{n!} |n\rangle \langle n| = E[e^{-\mu_2}] |0\rangle \langle 0| + E[e^{-\mu_2} \mu_2] |1\rangle \langle 1| + E[e^{-\mu_2} \mu_2^2] \omega_2 \rho_2 + \omega'_3 \rho_3, \quad (164)$$

where

$$\rho_2 := \frac{1}{\omega_2} \sum_{n=2}^{\infty} \frac{E[e^{-\mu_1} \mu_1^n]}{n! E[e^{-\mu_1} \mu_1^2]} |n\rangle \langle n| \quad (165)$$

$$\rho_3 := \frac{1}{\omega'_3} \sum_{n=3}^{\infty} \frac{E[e^{-\mu_2} \mu_2^n] E[e^{-\mu_1} \mu_1^2] - E[e^{-\mu_1} \mu_1^n] E[e^{-\mu_2} \mu_2^2]}{n! E[e^{-\mu_1} \mu_1^2]} |n\rangle \langle n| \quad (166)$$

$$\omega_2 := \sum_{n=2}^{\infty} \frac{E[e^{-\mu_1} \mu_1^n]}{n! E[e^{-\mu_1} \mu_1^2]} \quad (167)$$

$$\omega'_3 := \sum_{n=3}^{\infty} \frac{E[e^{-\mu_2} \mu_2^n] E[e^{-\mu_1} \mu_1^2] - E[e^{-\mu_1} \mu_1^n] E[e^{-\mu_2} \mu_2^2]}{n! E[e^{-\mu_1} \mu_1^2]}. \quad (168)$$

Indeed, our analysis in the previous sections uses the expansions (163) and (164) and their coefficients. Hence, we can apply the discussion with suitable modifications in the following way. (A similar idea is used in Wang [13, 14].)

In the following, we treat the case when the signal intensity is μ_1 and the decoy intensity is μ_2 . Then, we obtain the same argument with the following replacement: We replace the definition of ρ_2 in the above way, and define the set Ω_1 in the following way.

$$N^0 \in [Y_1^-(N, E[e^{-\mu_1}]), 2^{-\beta-6}), Y_1^+(N, E[e^{-\mu_1}]), 2^{-\beta-6})] \quad (169)$$

$$N^1 \in [Y_1^-(N, E[\mu_1 e^{-\mu_1}]), 2^{-\beta-6}), Y_1^+(N, E[\mu_1 e^{-\mu_1}]), 2^{-\beta-6})] \quad (170)$$

$$N_1^0 \in [Y_1^-(N_1, E[e^{-\mu_1}]), 2^{-\beta-6}), Y_1^+(N_1, E[e^{-\mu_1}]), 2^{-\beta-6})]$$

$$N_1^1 \in [Y_1^-(N_1, E[\mu_1 e^{-\mu_1}]), 2^{-\beta-6}), Y_1^+(N_1, E[\mu_1 e^{-\mu_1}]), 2^{-\beta-6})]$$

$$N_2^0 \in [Y_1^-(N_2, E[e^{-\mu_2}]), 2^{-\beta-6}), Y_1^+(N_2, E[e^{-\mu_2}]), 2^{-\beta-6})]$$

$$N_2^1 \in [Y_1^-(N_2, E[\mu_2 e^{-\mu_2}]), 2^{-\beta-6}), Y_1^+(N_2, E[\mu_2 e^{-\mu_2}]), 2^{-\beta-6})]$$

$$N_2^2 \in [Y_1^-(N_2, E[e^{-\mu_2} \mu_2^2] \omega_2), 2^{-\beta-6}), Y_1^+(N_2, E[e^{-\mu_2} \mu_2^2] \omega_2), 2^{-\beta-6})].$$

We also define \hat{N}_1 and \hat{N}_2 in the following way.

$$\hat{N}_1^0 := Y_1^-(N_1, E[e^{-\mu_1}]), 2^{-\beta-6})$$

$$\hat{N}_1^1 := Y_1^-(N_1, E[\mu_1 e^{-\mu_1}]), 2^{-\beta-6})$$

$$\hat{N}_2^0 := Y_1^-(N_2, E[e^{-\mu_2}]), 2^{-\beta-6})$$

$$\hat{N}_2^1 := Y_1^-(N_2, E[\mu_2 e^{-\mu_2}]), 2^{-\beta-6})$$

$$\hat{N}_2^2 := Y_1^-(N_2, E[e^{-\mu_2} \mu_2^2] \omega_2), 2^{-\beta-6}).$$

Then, we define \hat{N} by

$$\hat{N}^0 := Y_1^-(N, E[e^{-\mu_1}]), 2^{-\beta-6})$$

$$\hat{N}^1 := Y_1^-(N, E[\mu_1 e^{-\mu_1}]), 2^{-\beta-6})$$

Under the above modification, we change Condition 5 as follows.

$$\begin{aligned}
& Y_1^-(N_1, E[\mu_1 e^{-\mu_1}], 2^{-\beta-6}) Y_1^-(N_2, \omega_2 E[\mu_2^2 e^{-\mu_2}], 2^{-\beta-6}) \\
& > (N_1 - Y_1^-(N_1, E[e^{-\mu_1}], 2^{-\beta-6}) - Y_1^-(N_1, E[\mu_1 e^{-\mu_1}], 2^{-\beta-6})) Y_1^+(N_2, E[\mu_2 e^{-\mu_2}], 2^{-\beta-6}), \\
& Y_1^-(N_2, \omega_2 E[\mu_2^2 e^{-\mu_2}], 2^{-\beta-6}) Y_1^-(N, E[e^{-\mu_1}], 2^{-\beta-6}) \\
& > (N_1 - Y_1^-(N_1, E[e^{-\mu_1}], 2^{-\beta-6}) - Y_1^-(N_1, E[\mu_1 e^{-\mu_1}], 2^{-\beta-6})) Y_1^+(N_2, E[e^{-\mu_2}], 2^{-\beta-6}), \\
& \frac{Y_1^-(N_2, \omega_2 E[\mu_2^2 e^{-\mu_2}], 2^{-\beta-6})}{N_1 - Y_1^-(N_1, E[e^{-\mu_1}], 2^{-\beta-6}) - Y_1^-(N_1, E[\mu_1 e^{-\mu_1}], 2^{-\beta-6})} + \frac{Y_1^-(N_2, E[e^{-\mu_2}], 2^{-\beta-6})}{Y_1^+(N_1, E[e^{-\mu_1}], 2^{-\beta-6})} \\
& > \frac{2Y_1^+(N_2, E[\mu_2 e^{-\mu_2}], 2^{-\beta-6})}{Y_1^-(N_1, E[\mu_1 e^{-\mu_1}], 2^{-\beta-6})}.
\end{aligned}$$

Condition 6 is redefined in the term of Ω_1 defined above. Then, we define $\hat{\phi}_4(\mathbf{M})$ by using (123), (124), and Condition 6. Finally, we define the sacrifice bit length $S(\mathbf{M})$ to be $\hat{\phi}_4(\mathbf{M}) + 2\beta + 5$. Then, the relation (42) holds. Denoting the final state by $\rho_{A,E}$, we obtain

$$\|\rho_{A,E} - \rho_{\text{ideal}}\|_1 \leq 2^{-\beta}. \quad (171)$$

Even when the signal intensity is μ_2 and the decoy intensity is μ_1 , the above arguments are still valid by modifying (169), (170), (171), and (171) in the following way and replacing N , N^0 , and N^1 by N' , $N^{0'}$, and $N^{1'}$.

$$N^{0'} \in [Y_1^-(N', E[e^{-\mu_2}], 2^{-\beta-6}), Y_1^+(N', E[e^{-\mu_2}], 2^{-\beta-6})] \quad (172)$$

$$N^{1'} \in [Y_1^-(N', E[\mu_2 e^{-\mu_2}], 2^{-\beta-6}), Y_1^+(N', E[\mu_2 e^{-\mu_2}], 2^{-\beta-6})] \quad (173)$$

$$\hat{N}^0 := Y_1^-(N', E[e^{-\mu_2}], 2^{-\beta-6})$$

$$\hat{N}^1 := Y_1^-(N', E[\mu_2 e^{-\mu_2}], 2^{-\beta-6}).$$

Next, we treat the case when there are several candidates for the distribution of μ_2 and μ_1 while μ_2 and μ_1 obey independent and identical distributions. The possible distributions is denoted by $P_{\theta,1}$ and $P_{\theta,2}$, and the expectation is written by E_θ . Then, we denote the set Ω_1 under the distribution P_θ by $\Omega_{1,\theta}$.

In this case, Conditions 5 and 6 are needed to be satisfied for any θ . Hence, Condition 5 is redefined as follows. That is, the following relations hold for any θ .

$$\begin{aligned}
& Y_1^-(N_1, E_\theta[\mu_1 e^{-\mu_1}], 2^{-\beta-6}) Y_1^-(N_2, \omega_{2|\theta} E_\theta[\mu_2^2 e^{-\mu_2}], 2^{-\beta-6}) \\
& > (N_1 - Y_1^-(N_1, E_\theta[e^{-\mu_1}], 2^{-\beta-6}) - Y_1^-(N_1, E_\theta[\mu_1 e^{-\mu_1}], 2^{-\beta-6})) Y_1^+(N_2, E_\theta[\mu_2 e^{-\mu_2}], 2^{-\beta-6}), \\
& Y_1^-(N_2, \omega_{2|\theta} E_\theta[\mu_2^2 e^{-\mu_2}], 2^{-\beta-6}) Y_1^-(N, E_\theta[e^{-\mu_1}], 2^{-\beta-6}) \\
& > (N_1 - Y_1^-(N_1, E_\theta[e^{-\mu_1}], 2^{-\beta-6}) - Y_1^-(N_1, E_\theta[\mu_1 e^{-\mu_1}], 2^{-\beta-6})) Y_1^+(N_2, E_\theta[e^{-\mu_2}], 2^{-\beta-6}), \\
& \frac{Y_1^-(N_2, \omega_{2|\theta} E_\theta[\mu_2^2 e^{-\mu_2}], 2^{-\beta-6})}{N_1 - Y_1^-(N_1, E_\theta[e^{-\mu_1}], 2^{-\beta-6}) - Y_1^-(N_1, E_\theta[\mu_1 e^{-\mu_1}], 2^{-\beta-6})} + \frac{Y_1^-(N_2, E_\theta[e^{-\mu_2}], 2^{-\beta-6})}{Y_1^+(N_1, E_\theta[e^{-\mu_1}], 2^{-\beta-6})} \\
& > \frac{2Y_1^+(N_2, E_\theta[\mu_2 e^{-\mu_2}], 2^{-\beta-6})}{Y_1^-(N_1, E_\theta[\mu_1 e^{-\mu_1}], 2^{-\beta-6})},
\end{aligned}$$

where $\omega_{2|\theta}$ is ω_2 with the distribution $P_{\theta,1}$.

Further, we redefine Condition 6 as the condition that all of C_1^0 , C_1^1 , C_2^0 , C_2^1 , and C_2^2 are negative for $\vec{N} \in \cup_\theta \Omega_{1,\theta}$. We define $\hat{\phi}_{4,\theta}(\mathbf{M})$ to be $\hat{\phi}_4(\mathbf{M})$ given in (123) and

(124) when the true distributions are $P_{\theta,1}$ and $P_{\theta,2}$. Finally, we define the sacrifice bit length $S(\mathbf{M})$ by $\sup_{\theta} \hat{\phi}_{4,\theta}(\mathbf{M}) + 2\beta + 5$. Then, letting $\rho_{A,E|\theta}$ be the final state with the true distributions $P_{\theta,1}$ and $P_{\theta,2}$, $\rho_{\text{ideal}|\theta}$ be the ideal state, we obtain

$$\|\rho_{A,E|\theta} - \rho_{\text{ideal}|\theta}\|_1 \leq 2^{-\beta}. \quad (174)$$

That is, the inequality holds for any θ .

In the following, we consider the case when the pulses are generated with the mixture of the plural independent and identical distributions $P_{\theta,1}$ and $P_{\theta,2}$, respectively. Then, the intensities of $N + N_1$ pulses are described by $(\mu_{1,1}, \dots, \mu_{1,N+N_1})$ and are subject to the distribution $\sum_{\theta} \lambda_{\theta} P_{\theta,1}^{\times(N+N_1)}$, where $P^{\times N}$ is the N -fold independent and identical distribution of P . Similarly, the intensities of $N' + N_2$ pulses are described by $(\mu_{2,1}, \dots, \mu_{2,N_2})$ and are subject to the distribution $\sum_{\theta} \lambda_{\theta} P_{\theta,2}^{\times N_2}$. Since the final state is $\sum_{\theta} \lambda_{\theta} \rho_{A,E|\theta}$, we obtain

$$\|(\sum_{\theta} \lambda_{\theta} \rho_{A,E|\theta}) - (\sum_{\theta} \lambda_{\theta} \rho_{\text{ideal}|\theta})\|_1 \leq \sum_{\theta} \lambda_{\theta} \|\rho_{A,E|\theta} - \rho_{\text{ideal}|\theta}\|_1 \leq 2^{-\beta}. \quad (175)$$

Hence, the universal composability is upperly bounded by $2^{-\beta}$.

12.2. Case when the distributions of intensities μ_1 are μ_2 are not fixed

Finally, we consider the case when the distribution of intensities μ_1 are μ_2 are changed during one coding block. For simplicity, we treat the case when there are two distributions. That is, the intensity μ_1 obeys the distribution $P_{\theta,1}$ or the other distribution $P_{\theta',1}$, and the other intensity μ_2 obeys the distribution $P_{\eta,2}$ or the other distribution $P_{\eta',2}$.

Then, we assume that the intensities $(\mu_{1,1}, \dots, \mu_{N_1})$ of N_1 pulses satisfy $P_{\theta,1}^{\times N_{1|\theta}} \times P_{\theta',1}^{\times N_{1|\theta'}}$, and the intensities $(\mu_{2,1}, \dots, \mu_{N_2})$ of N_2 pulses satisfy $P_{\eta,2}^{\times N_{2|\eta}} \times P_{\eta',2}^{\times N_{2|\eta'}}$, where $N_1 = N_{1|\theta} + N_{1|\theta'}$ and $N_2 = N_{2|\eta} + N_{2|\eta'}$. We denote the state ρ_2 defined in (165) and the real number ω_2 defined in (167) under the distribution ω_2 by $\rho_{2|\theta}$ and $\omega_{2|\theta}$, respectively. We also denote \bar{q}_{\times}^2 and \bar{b}_{\times}^2 under the distribution ω_2 by $\bar{q}^{2|\theta}$ and $\bar{b}^{2|\theta}$, respectively. Then, we denote the breakdown of $N_{1|\theta}$ pulses by $(N_{1|\theta}^0, N_{1|\theta}^1, N_{1|\theta}^2)$. We make the same definitions for θ' .

Defining $N_1^0 := N_{1|\theta}^0 + N_{1|\theta'}^0$, $N_1^1 := N_{1|\theta}^1 + N_{1|\theta'}^1$, $N_1^2 := N_{1|\theta}^2 + N_{1|\theta'}^2$, $\bar{q}_{\times}^2 := \frac{\bar{q}_{\times|\theta}^2 N_{1|\theta}^2 + \bar{q}_{\times|\theta'}^2 N_{1|\theta'}^2}{N_1^2}$, $\bar{b}_{\times}^2 := \frac{\bar{b}_{\times|\theta}^2 N_{1|\theta}^2 + \bar{b}_{\times|\theta'}^2 N_{1|\theta'}^2}{N_1^2}$, and $\rho_{2,\times} := \frac{\rho_{2,\times|\theta} N_{1|\theta}^2 + \rho_{2,\times|\theta'} N_{1|\theta'}^2}{N_1^2}$, we make the following expansion of the pulse subject to the distribution $P_{\eta,2}$.

$$\begin{aligned} \sum_{n=0}^{\infty} \frac{\mathbb{E}_{\eta}[e^{-\mu_2} \mu_2^n]}{n!} |n\rangle \langle n| &= \mathbb{E}_{\eta}[e^{-\mu_2}] |0\rangle \langle 0| + \mathbb{E}_{\eta}[e^{-\mu_2} \mu_2] |1\rangle \langle 1| + \mathbb{E}_{\eta}[e^{-\mu_2} \mu_2^2] \omega_{2|(N_{1|\theta}^2, N_{1|\theta'}^2)} \rho_2 \\ &\quad + \omega'_{3|(N_{1|\theta}^2, N_{1|\theta'}^2), \eta} \rho_{3|\eta}, \end{aligned} \quad (176)$$

where

$$\begin{aligned}\rho_{3|\eta} &:= \frac{1}{\omega'_3} \sum_{n=3}^{\infty} \left(E_{\eta}[e^{-\mu_2} \mu_2^n] - \frac{E_{\eta}[e^{-\mu_2} \mu_2^2] \frac{\frac{E_{\theta}[e^{-\mu_1} \mu_1^n]}{\omega_{2,\times|\theta}} N_{1|\theta}^2 + \frac{E_{\theta'}[e^{-\mu_1} \mu_1^n]}{\omega_{2,\times|\theta'}} N_{1|\theta'}^2}{\frac{N_{1|\theta}^2}{\omega_{2,\times|\theta}} + \frac{N_{1|\theta'}^2}{\omega_{2,\times|\theta'}}}} \right) |n\rangle \langle n| \\ &= \frac{1}{\omega'_3} \sum_{n=3}^{\infty} \left(E_{\eta}[e^{-\mu_2} \mu_2^n] - \frac{E_{\eta}[e^{-\mu_2} \mu_2^2] \frac{E_{\theta}[e^{-\mu_1} \mu_1^n]}{\omega_{2,\times|\theta}} N_{1|\theta}^2 + \frac{E_{\theta'}[e^{-\mu_1} \mu_1^n]}{\omega_{2,\times|\theta'}} N_{1|\theta'}^2}{\frac{N_{1|\theta}^2}{\omega_{2,\times|\theta}} + \frac{N_{1|\theta'}^2}{\omega_{2,\times|\theta'}}}} \right) |n\rangle \langle n| \end{aligned} \quad (177)$$

and

$$\omega_{2|(N_{1|\theta}^2, N_{1|\theta'}^2)} := \left(\frac{\frac{N_{1|\theta}^2}{\omega_{2,\times|\theta}} + \frac{N_{1|\theta'}^2}{\omega_{2,\times|\theta'}}}{N_1^2} \right)^{-1} \in [\omega_{2,\times|\min}, \omega_{2,\times|\max}] \quad (178)$$

$$\omega'_{3|(N_{1|\theta}^2, N_{1|\theta'}^2), \eta} := \sum_{n=3}^{\infty} \left(E_{\eta}[e^{-\mu_2} \mu_2^n] - \frac{E_{\eta}[e^{-\mu_2} \mu_2^2] \frac{E_{\theta}[e^{-\mu_1} \mu_1^n]}{\omega_{2,\times|\theta}} N_{1|\theta}^2 + \frac{E_{\theta'}[e^{-\mu_1} \mu_1^n]}{\omega_{2,\times|\theta'}} N_{1|\theta'}^2}{\frac{N_{1|\theta}^2}{\omega_{2,\times|\theta}} + \frac{N_{1|\theta'}^2}{\omega_{2,\times|\theta'}}}} \right) \quad (179)$$

$$\omega_{2,\times|\min} := \min(\omega_{2,\times|\theta}, \omega_{2,\times|\theta'}), \quad \omega_{2,\times|\max} := \max(\omega_{2,\times|\theta}, \omega_{2,\times|\theta'}). \quad (180)$$

Next, we denote the breakdown of $N_{2|\eta}$ pulses by $(N_{2|\eta}^0, N_{2|\eta}^1, N_{2|\eta}^2, N_{2|\eta}^3)$. We also denote \bar{q}^3 and \bar{b}_{\times}^3 under the distribution $P_{\eta,2}$ by $\bar{q}^{3|\eta}$ and $\bar{b}^{3|\eta}$, respectively. We make the same definitions for η' .

Defining $N_2^0 := N_{2|\eta}^0 + N_{2|\eta'}^0$, $N_2^1 := N_{2|\eta}^1 + N_{2|\eta'}^1$, $N_2^2 := N_{2|\eta}^2 + N_{2|\eta'}^2$, and $N_2^3 := N_{2|\eta}^3 + N_{2|\eta'}^3$, we obtain the following relations by the same way as the relations (65), (66), (67), and (74).

$$\begin{aligned}\bar{M}_0 &= \bar{q}^0 N_0 \\ \bar{M}_1 &= \frac{\bar{q}^0}{2} N_1^0 + \bar{a}_{\times}^1 N_1^1 + \bar{a}_{\times}^2 N_1^2 \\ \bar{M}_2 &= \frac{\bar{q}^0}{2} N_2^0 + \bar{a}_{\times}^1 N_2^1 + \bar{a}_{\times}^2 N_2^2 + \bar{a}_{\times|\eta}^3 N_{2|\eta}^3 + \bar{a}_{\times|\eta'}^3 N_{2|\eta'}^3 \\ \bar{M}_3 &= \frac{\bar{q}^0}{2} N_1^0 + \bar{b}_{\times}^1 N_1^1 + \bar{b}_{\times}^2 N_1^2. \end{aligned} \quad (181)$$

Hence, using the same way as Subsection 7.2, we obtain the upper bound of ϕ by solving the above equations with $\bar{p}_{3,\times|\eta} = \bar{p}_{3,\times|\eta'} = \bar{b}_{\times|\eta}^3 = \bar{b}_{\times|\eta'}^3 = 0$. That is, defining

$$\begin{aligned}\hat{q}^0(\bar{M}_0) &:= \frac{\bar{M}_0}{N_0} \\ \hat{a}_{\times}^1(\mathbf{M}, \vec{N}) &:= \frac{N_2^2(\bar{M}_1 - \bar{M}_0 N_1^0/2N_0) - N_1^2(\bar{M}_2 - \bar{M}_0 N_2^0/2N_0)}{N_1^1 N_2^2 - N_2^1 N_1^2} \\ \hat{b}_{\times}^1(\mathbf{M}, \vec{N}) &:= \frac{\bar{M}_3 - \bar{M}_0 N_1^0/2N_0}{N_1^1}, \end{aligned}$$

we define

$$\hat{\phi}_3(\bar{\mathbf{M}}, \vec{N}) := \hat{\phi}_1(\mathbf{N}, \hat{q}^0(\bar{M}_0), \hat{a}_{\times}^1(\mathbf{M}, \vec{N}), \hat{b}_{\times}^1(\mathbf{M}, \vec{N})).$$

Hence, the same arguments as the previous discussion hold with the following replacement.

In the following, we treat the case when the signal intensity is μ_1 and the decoy intensity is μ_2 . Then, the set Ω_1 is redefined as the set of \vec{N} satisfying the following conditions.

$$N^0 \in [\min_{\theta} Y_1^-(N, E_{\theta}[e^{-\mu_1}], 2^{-\beta-6}), \max_{\theta} Y_1^+(N, E_{\theta}[e^{-\mu_1}], 2^{-\beta-6})] \quad (182)$$

$$N^1 \in [\min_{\theta} Y_1^-(N, E_{\theta}[\mu_1 e^{-\mu_1}], 2^{-\beta-6}), \max_{\theta} Y_1^+(N, E_{\theta}[\mu_1 e^{-\mu_1}], 2^{-\beta-6})] \quad (183)$$

$$N_1^0 \in [\min_{\theta} Y_1^-(N_1, E_{\theta}[e^{-\mu_1}], 2^{-\beta-6}), \max_{\theta} Y_1^+(N_1, E_{\theta}[e^{-\mu_1}], 2^{-\beta-6})]$$

$$N_1^1 \in [\min_{\theta} Y_1^-(N_1, E_{\theta}[\mu_1 e^{-\mu_1}], 2^{-\beta-6}), \max_{\theta} Y_1^+(N_1, E_{\theta}[\mu_1 e^{-\mu_1}], 2^{-\beta-6})]$$

$$N_2^0 \in [\min_{\eta} Y_1^-(N_2, E_{\eta}[e^{-\mu_2}], 2^{-\beta-6}), \max_{\eta} Y_1^+(N_2, E_{\eta}[e^{-\mu_2}], 2^{-\beta-6})]$$

$$N_2^1 \in [\min_{\eta} Y_1^-(N_2, E_{\eta}[\mu_2 e^{-\mu_2}], 2^{-\beta-6}), \max_{\eta} Y_1^+(N_2, E_{\eta}[\mu_2 e^{-\mu_2}], 2^{-\beta-6})]$$

$$N_2^2 \in [\min_{\eta} Y_1^-(N_2, E_{\eta}[e^{-\mu_2} \mu_2^2] \omega_{2,\times|\min}, 2^{-\beta-6}), \max_{\eta} Y_1^+(N_2, E_{\eta}[e^{-\mu_2} \mu_2^2] \omega_{2,\times|\max}, 2^{-\beta-6})].$$

Then, the relation $\Pr(\vec{N} \in \Omega_1^c) \leq 14 \cdot 2^{-\beta-6}$ holds. We define \hat{N}_1 and \hat{N}_2 as

$$\hat{N}_1^0 := \min_{\theta} Y_1^-(N_1, E_{\theta}[e^{-\mu_1}], 2^{-\beta-6})$$

$$\hat{N}_1^1 := \min_{\theta} Y_1^-(N_1, E_{\theta}[\mu_1 e^{-\mu_1}], 2^{-\beta-6})$$

$$\hat{N}_2^0 := \min_{\eta} Y_1^-(N_2, E_{\eta}[e^{-\mu_2}], 2^{-\beta-6})$$

$$\hat{N}_2^1 := \min_{\eta} Y_1^-(N_2, E_{\eta}[\mu_2 e^{-\mu_2}], 2^{-\beta-6})$$

$$\hat{N}_2^2 := \min_{\eta} Y_1^-(N_2, E_{\eta}[e^{-\mu_2} \mu_2^2] \omega_{2,\times|\min}, 2^{-\beta-6}),$$

and \hat{N} as

$$\hat{N}^0 := \min_{\theta} Y_1^-(N, E_{\theta}[e^{-\mu_1}], 2^{-\beta-6}) \quad (184)$$

$$\hat{N}^1 := \min_{\theta} Y_1^-(N, E_{\theta}[\mu_1 e^{-\mu_1}], 2^{-\beta-6}). \quad (185)$$

Then, Condition 5 is modified to

$$\begin{aligned} & \min_{\theta} Y_1^-(N_1, E_{\theta}[\mu_1 e^{-\mu_1}], 2^{-\beta-6}) \min_{\eta} Y_1^-(N_2, \omega_{2,\times|\min} E_{\eta}[\mu_2^2 e^{-\mu_2}], 2^{-\beta-6}) \\ & > \max_{\theta} (N_1 - Y_1^-(N_1, E_{\theta}[e^{-\mu_1}], 2^{-\beta-6}) - Y_1^-(N_1, E_{\theta}[\mu_1 e^{-\mu_1}], 2^{-\beta-6})) \max_{\eta} Y_1^+(N_2, E_{\eta}[\mu_2 e^{-\mu_2}], 2^{-\beta-6}), \\ & \min_{\eta} Y_1^-(N_2, \omega_{2,\times|\min} E_{\eta}[\mu_2^2 e^{-\mu_2}], 2^{-\beta-6}) \min_{\theta} Y_1^-(N, E_{\theta}[e^{-\mu_1}], 2^{-\beta-6}) \\ & > \max_{\theta} (N_1 - Y_1^-(N_1, E_{\theta}[e^{-\mu_1}], 2^{-\beta-6}) - Y_1^-(N_1, E_{\theta}[\mu_1 e^{-\mu_1}], 2^{-\beta-6})) \max_{\eta} Y_1^+(N_2, E_{\eta}[e^{-\mu_2}], 2^{-\beta-6}), \\ & \frac{\min_{\eta} Y_1^-(N_2, \omega_{2,\times|\min} E_{\eta}[\mu_2^2 e^{-\mu_2}], 2^{-\beta-6})}{\max_{\theta} N_1 - Y_1^-(N_1, E_{\theta}[e^{-\mu_1}], 2^{-\beta-6}) - Y_1^-(N_1, E_{\theta}[\mu_1 e^{-\mu_1}], 2^{-\beta-6})} + \frac{\min_{\eta} Y_1^-(N_2, E_{\eta}[e^{-\mu_2}], 2^{-\beta-6})}{\max_{\theta} Y_1^+(N_1, E_{\theta}[e^{-\mu_1}], 2^{-\beta-6})} \\ & > \frac{2 \max_{\eta} Y_1^+(N_2, E_{\eta}[\mu_2 e^{-\mu_2}], 2^{-\beta-6})}{\min_{\theta} Y_1^-(N_1, E_{\theta}[\mu_1 e^{-\mu_1}], 2^{-\beta-6})}. \end{aligned}$$

When Condition 5 holds under this modification, Conditions 1 and 2 hold with any $\vec{N} \in \Omega_1$. Then, we define $\hat{\phi}_4(\mathbf{M})$ by using (123), (124), and Condition 6 concerning the above defined set Ω_1 . Defining the sacrifice bit length $S(\mathbf{M})$ as $\hat{\phi}_4(\mathbf{M}) + 2\beta + 5$, we obtain (42). The final state $\rho_{A,E}$ satisfies

$$\|\rho_{A,E|\theta} - \rho_{\text{ideal}|\theta}\|_1 \leq 2^{-\beta}. \quad (186)$$

That is, we derive the sacrifice bit length $S(\mathbf{M})$ whose security is guaranteed.

Even when the signal intensity is μ_2 and the decoy intensity is μ_1 , the above argument are still valid by modifying (182), (183), (184), and (185) in the following way and replacing N , N^0 , and N^1 by N' , $N^{0'}$, and $N^{1'}$.

$$N^{0'} \in [\min_{\eta} Y_1^-(N', E_{\eta}[e^{-\mu_2}], 2^{-\beta-6}), \max_{\eta} Y_1^+(N', E_{\eta}[e^{-\mu_2}], 2^{-\beta-6})] \quad (187)$$

$$N^{1'} \in [\min_{\eta} Y_1^-(N', E_{\eta}[\mu_2 e^{-\mu_2}], 2^{-\beta-6}), \max_{\eta} Y_1^+(N', E_{\eta}[\mu_2 e^{-\mu_2}], 2^{-\beta-6})] \quad (188)$$

$$\hat{N}^{0'} := \min_{\eta} Y_1^-(N', E_{\eta}[e^{-\mu_2}], 2^{-\beta-6})$$

$$\hat{N}^{1'} := \min_{\eta} Y_1^-(N', E_{\eta}[\mu_2 e^{-\mu_2}], 2^{-\beta-6}).$$

When there are more than 3 candidates for the distributions P_1 and P_2 , generalizing the definition of Ω_1 and Condition 5, we can define $\hat{\phi}_4(\mathbf{M})$ and the sacrifice bit length $S(\mathbf{M}) := \hat{\phi}_4(\mathbf{M}) + 2\beta + 5$. Then, the relation (186) holds.

13. Numerical comparison

Finally, we treat numerical analysis with the finite-block length. In the following, we consider only the case when the signal intensity is μ_2 , the decoy intensity is μ_1 , $N_0 = N_1 = N_2 = N'/10$, and $\beta = 80$, i.e., the trace norm is less than 2^{-80} .

13.1. Fixed intensities

First, we treat the case with two fixed intensities μ_1 and μ_2 . It is natural to assume that the measured values M_0 , M_1 , M_2 , and M_3 are given as

$$M_0 = p_0 N_0, \quad M_1 = (p_{1,\times} - s_{1,\times}) N_1 \quad (189)$$

$$M_2 = (p_{2,\times} - s_{2,\times}) N_2, \quad M_3 = s_{1,\times} N_1 \quad (190)$$

with

$$p_{i,+} = p_{i,\times} = 1 - e^{-\alpha\mu_i} + p_0, \quad s_{i,+} = s_{i,\times} = s(1 - e^{-\alpha\mu_i}) + \frac{p_0}{2}. \quad (191)$$

We choose N' to be $M'/p_{2,+}$. Then, we consider the key generation rate

$$R := \frac{M' - S - \eta h\left(\frac{s_{1,+}}{p_{1,+}}\right)M}{N'}, \quad (192)$$

where S is the sacrifice bit length given in Section 10.

As is illustrated in Figs. 10 and 11 with $\mu_1 = 0.1$, $\alpha = 1/1000$, $p_0 = 0.0000004$, $\eta = 1.1$, the key generation rate is close to the asymptotic key generation rate $R_2(\mu_1, \mu_2)$ when the length of the code M is increasing.

We show that the asymptotic key generation rate $R_2(\mu_1, \mu_2)$ is monotone decreasing concerning μ_1 in Theorem 3. However, as is illustrated in Fig. 12 with $\alpha = 1/1000$, $p_0 = 0.0000004$, $\eta = 1.1$, the key generation rate is not monotone decreasing concerning μ_1 when the length of the code M is not sufficiently large. That is, too small μ_1 does not give a good key generation rate. This is because smaller μ_1 yields a larger estimation error.

Substituting the approximation formula for the sacrifice bit length (162) into (192), we obtain an approximation rate \tilde{R} of the key generation rate R . As is illustrated in Fig. 13 with $\mu_2 = 0.5$, $\alpha = 1/1000$, $p_0 = 0.0000004$, $\eta = 1.1$, the approximation rate \tilde{R} is close to the rate R when the length of the code M is sufficiently large. However, as is illustrated in Fig. 14 with $\mu_2 = 0.5$, $\alpha = 1/1000$, $p_0 = 0.0000004$, $\eta = 1.1$, when the length of the code M is not sufficiently large, the approximation rate \tilde{R} is not so close to the rate R , but the shape of graph of \tilde{R} is similar to that of R .

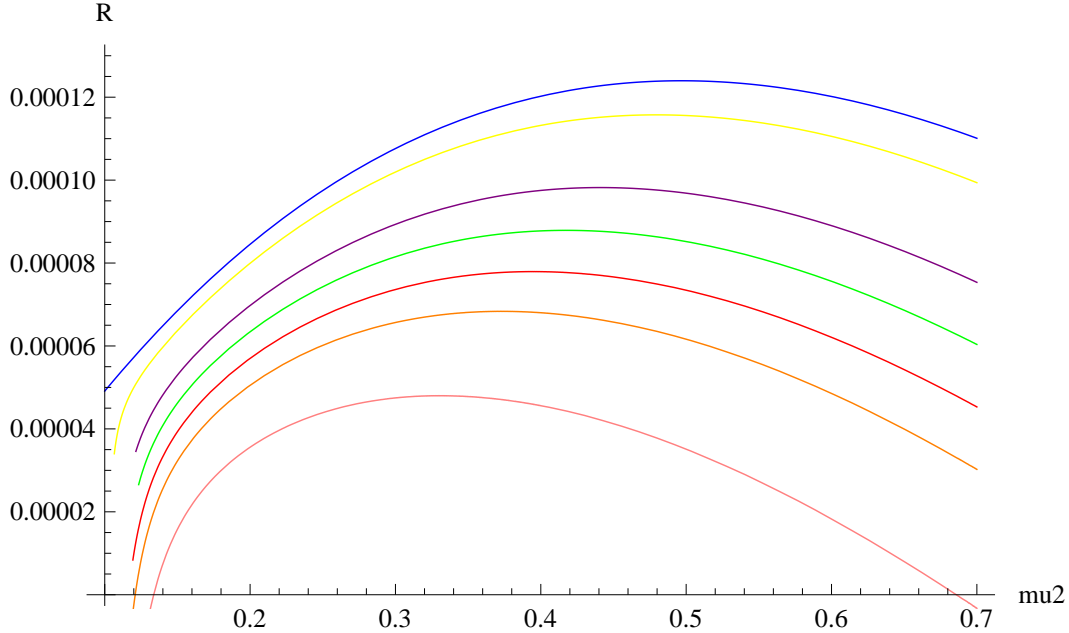


Figure 10. The above graphs describes the rate R given in (192) as functions of μ_2 when $\mu_1 = 0.1$. The pink line is the case with $M = 10^6$. The orange line is the case with $M = 2 \times 10^6$. The red line is the case with $M = 3 \times 10^6$. The green line is the case with $M = 5 \times 10^6$. The purple line is the case with $M = 10^7$. The yellow line is the case with $M = 10^8$. The blue line is the asymptotic case.

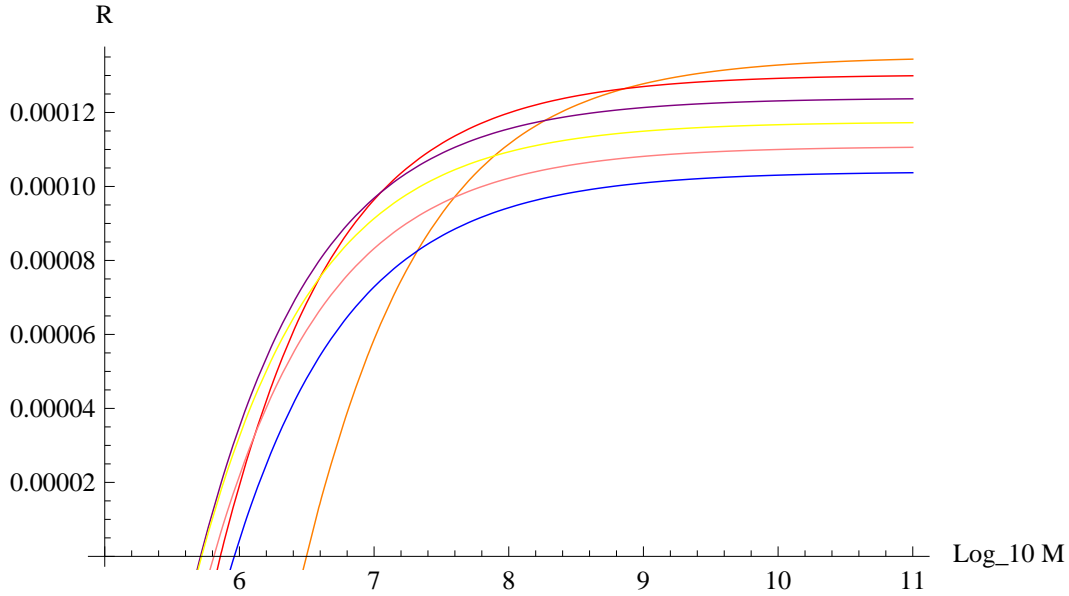


Figure 11. The above graphs describes the rate R given in (192) as functions of M when $\mu_2 = 0.5$. The orange line is the case with $\mu_1 = 0.01$. The red line is the case with $\mu_1 = 0.05$. The purple line is the case with $\mu_1 = 0.1$. The yellow line is the case with $\mu_1 = 0.15$. The pink line is the case with $\mu_1 = 0.2$. The blue line is the case with $\mu_1 = 0.25$.

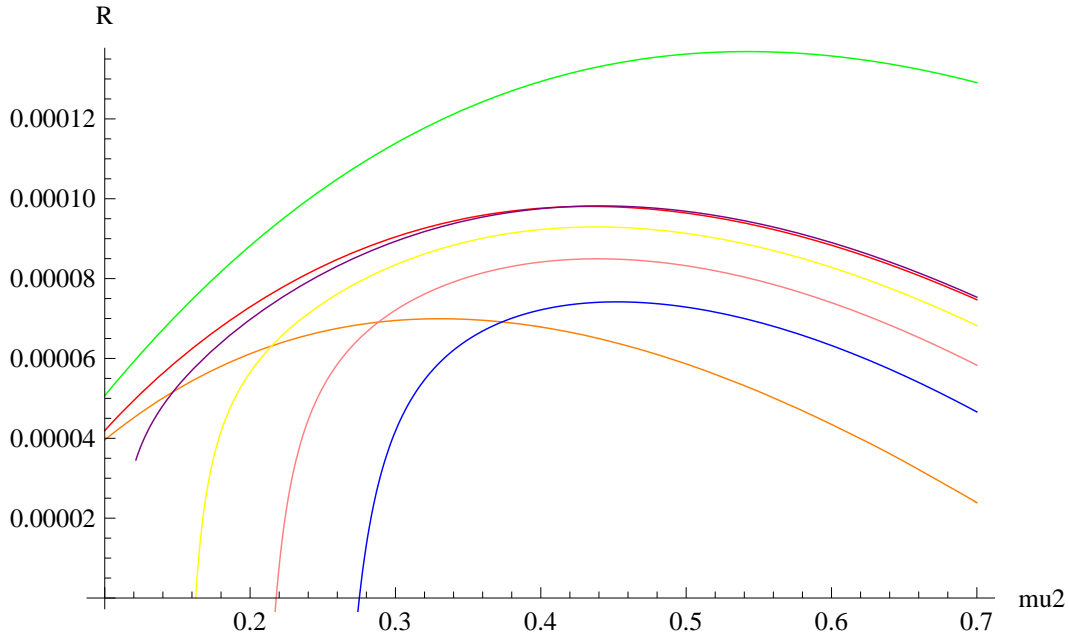


Figure 12. The above graphs describes the rate R given in (192) as functions of μ_2 when $M = 10^7$. The orange line is the case with $\mu_1 = 0.01$. The red line is the case with $\mu_1 = 0.05$. The purple line is the case with $\mu_1 = 0.1$. The yellow line is the case with $\mu_1 = 0.15$. The pink line is the case with $\mu_1 = 0.2$. The blue line is the case with $\mu_1 = 0.25$. The green line is the asymptotic case with $\mu_1 \rightarrow 0$.

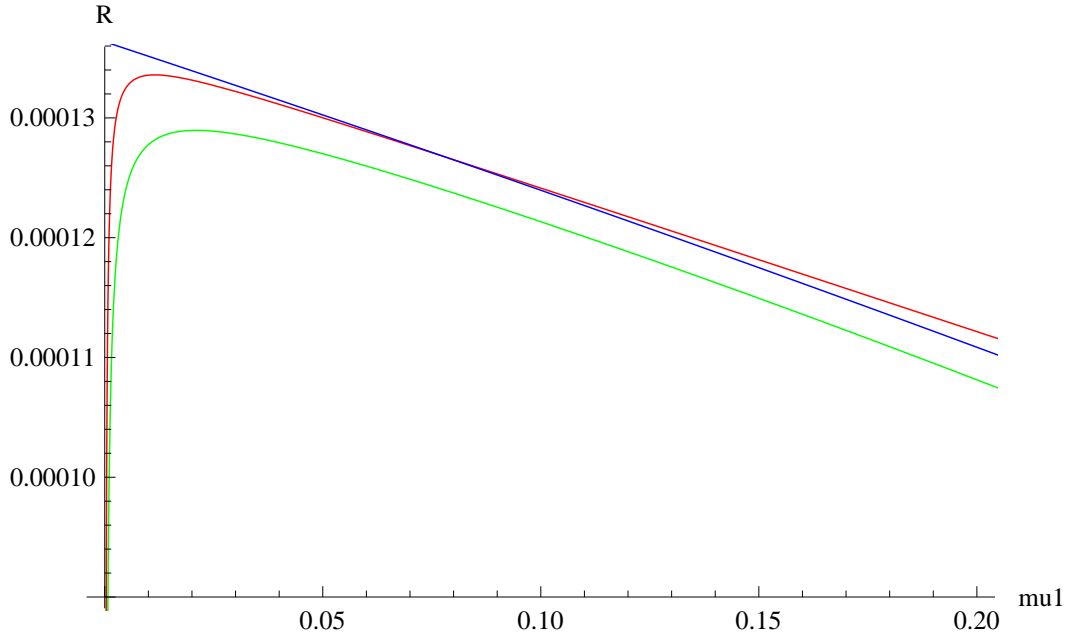


Figure 13. The vertical axis describes the rate. The horizontal axis describes the decoy intensity μ_1 . All graphs give the rates with $M = 10^9$ and $\mu_2 = 0.5$. The green line is the rate R given in (192). The red line is the approximation rate \tilde{R} . The blue line is the approximation rate \tilde{R} .

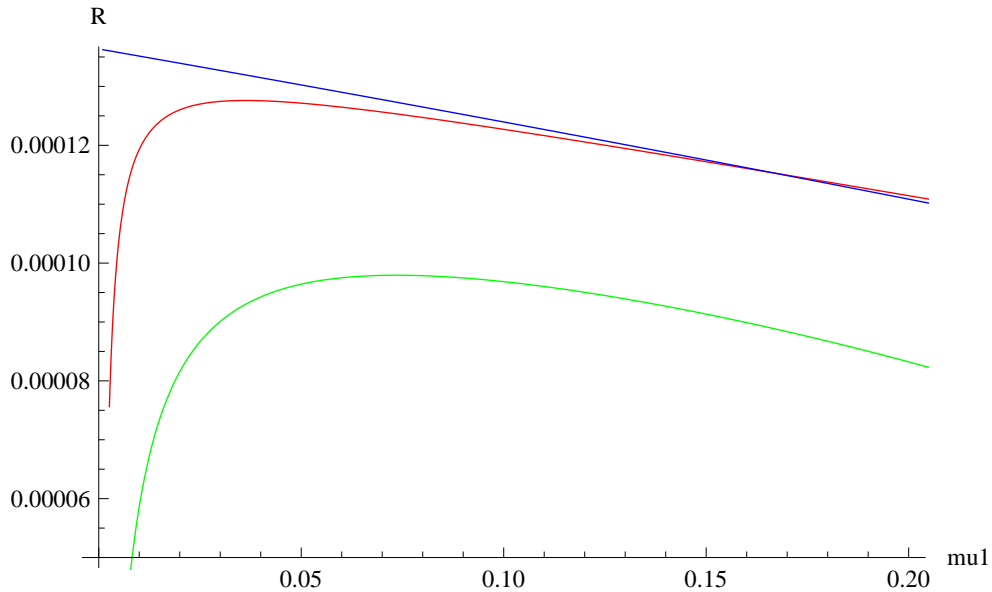


Figure 14. The vertical axis describes the rate. The horizontal axis describes the decoy intensity μ_1 . All graphs give the rates with $M = 10^7$ and $\mu_2 = 0.5$. The green line is the rate R given in (192). The red line is the approximation rate \tilde{R} . The blue line is the approximation rate \tilde{R} .

13.2. Gaussian distribution

Next, we treat the case when two intensities μ_1 and μ_2 obey the Gaussian distributions with the averages $\bar{\mu}_1$ and $\bar{\mu}_2$ and the standard deviations $\bar{\mu}_1 t$ and $\bar{\mu}_2 t$, respectively[].

In order to calculate the sacrifice bit length given in Subsection 12.1, we need $E[e^{\mu_i}]$, $E[\mu_i e^{\mu_i}]$, $E[\mu_i^2 e^{\mu_i}]$, and ω_2 , which can be easily calculate from the calculations given in Appendix D. In this case, due to (D.6), it is natural that $p_{i,+}$, $p_{i,\times}$, $s_{i,+}$, and $s_{i,\times}$ are given as

$$p_{i,+} = p_{i,\times} = 1 - E[e^{-\alpha\mu_i}] + p_0 = 1 - e^{\frac{2\alpha\bar{\mu}_i - \alpha^2 t^2 \bar{\mu}_1^2}{2}} + p_0 \quad (193)$$

$$s_{i,+} = s_{i,\times} = s(1 - E[e^{-\alpha\mu_i}]) + \frac{p_0}{2} = s(1 - e^{\frac{2\alpha\bar{\mu}_i - \alpha^2 t^2 \bar{\mu}_1^2}{2}}) + \frac{p_0}{2}. \quad (194)$$

Hence, the measured values M_0 , M_1 , M_2 , and M_3 are given by (189) and (190) with the above $p_{i,\times}$ and $s_{i,\times}$. We choose N' to be $M'/p_{2,+}$. Then, substituting the sacrifice bit length given in Subsection 12.1 into the key generation rate R (192), we obtain the numerical calculation in Figs. 15 and 16.

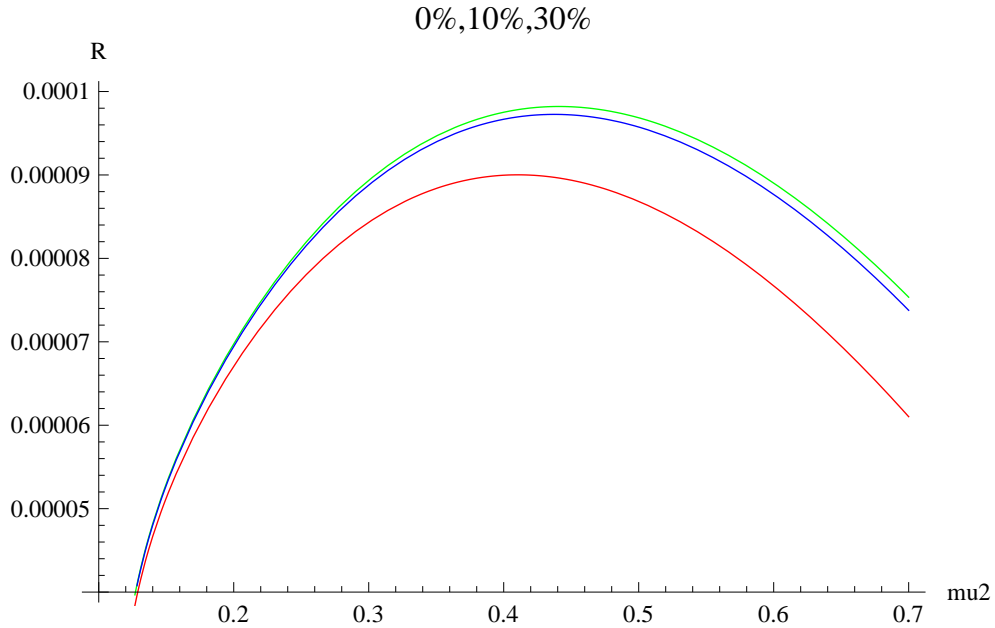


Figure 15. The vertical axis describes the rate R . The horizontal axis describes the signal intensity μ_2 . All graphs give the rates R with $M = 10^7$ and $\mu_1 = 0.1$. The green line is the rate R with $t = 0\%$. The blue line is the rate R with $t = 10\%$. The red line is the rate R with $t = 30\%$.

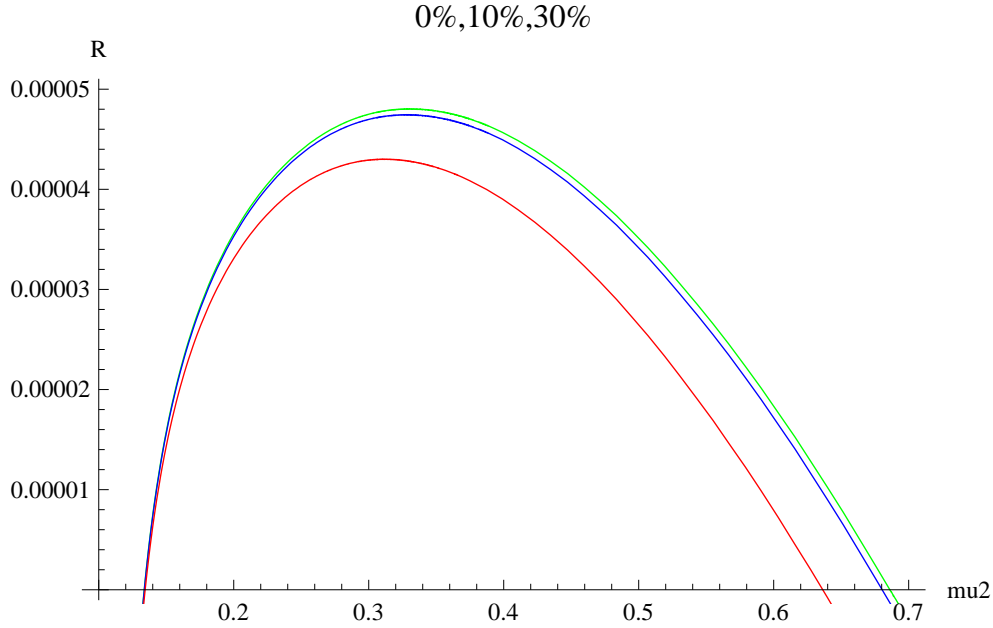


Figure 16. The vertical axis describes the rate R . The horizontal axis describes the signal intensity μ_2 . All graphs give the rates R with $M = 10^6$ and $\mu_1 = 0.1$. The green line is the rate R with $t = 0\%$. The blue line is the rate R with $t = 10\%$. The red line is the rate R with $t = 30\%$.

14. Conclusion and further improvement

In this paper, in the BB84 protocol with the decoy method, based on several observed values, we have derived the required sacrifice bit-length $S(\mathbf{M}) = \hat{\phi}_4(\mathbf{M}) + 2\beta + 5$, where $\hat{\phi}_4(\mathbf{M})$ is given in (123) and (124). Under the above sacrifice bit-length, we have shown that the final keys satisfy the security condition $\|\rho_{A,E} - \rho_{\text{ideal}}\|_1 \leq 2^{-\beta}$ when the parameters μ_1 , μ_2 , N , N_0 , N_1 , and N_2 satisfy Condition 5. Hence, in order to apply our formula, we need to choose the parameters μ_1 , μ_2 , N , N_0 , N_1 , and N_2 so that Condition 5 holds. This is a definitive requirement for our analysis. However, when we choose sufficiently large integers N , N_0 , N_1 , and N_2 for the two values μ_1 and $\mu_2 - \mu_1$, Condition 5 holds. Indeed, when the two positive values μ_1 and $\mu_2 - \mu_1$ are quite small, we need to choose quite large integers N , N_0 , N_1 , and N_2 . As the second requirement, we need to choose the parameters μ_1 , μ_2 , N , N_0 , N_1 , and N_2 so that Conditions 4 and 6 hold with a high probability when there is no eavesdropper. This requirement is also satisfied when the integers N , N_0 , N_1 , and N_2 are sufficiently large and the noise in the channel is sufficiently small.

Since the decoy method has so many parameters, it is quite difficult to derived tight evaluation. The proposed method might be improved by modifying several points. However, such a modification might make the protocol complex. For example, while we treat the decoding phase error probability and the estimation error probability, separately, Hayashi and Tsurumaru [32] treated them jointly. In order to keep the simplicity, it is better to treat these terms separately. Further, in Section 5, we proposed

to treat the probability based on the hyper geometric distribution by using the binomial distribution. If we treat the probabilities given in Section 7 with the hyper geometric distribution, we obtain a better evaluation, but our analysis becomes much harder.

In order to treat this problem, we have to consider the trade-off between the complexity and the tightness of our evaluation. This kind of trade-off cannot be ignored from an industrial view point. If the protocol is more complex, the cost for maintenance becomes higher. In particular, when we change the arrangement of the total system or we change the parameter of the system, we have to rewrite the program for calculating the sacrifice bit-length. If the protocol is simple, the change can be easily done. Otherwise, it spends some additional cost. Hence, we have to take into account this trade-off. Due to the property, our treatment is heuristic.

However, its systematic treatment might be possible partially in the following sense. Assume that we employ the Renner's formalism. If we parametrize the channel with more parameters to be estimated, the asymptotic key generation rate becomes better. One might consider that if the number of parameters describing the model increases, we obtain a better estimation of the model. However, it is considered that it is not true in statistics. This is because if we do not have enough data to characterize so many parameters, we obtain a larger error. In order to resolve this problem, we have to treat the trade-off between the error and the number of parameters. Such a problem is called the model selection. In order to treat this problem quantitatively, we can use several information criteria, Akaike information criterion (AIC)[41], Takeuchi information criterion (TIC)[42], and minimum description length principle (MDL)[43]. If we employ the Renner's formalism, and increase the number of channel parameters for precise description of channel, we need to consider this kind of trade-off. Currently, it is not known that what kind of information criterion is suitable for the above our trade-off.

Acknowledgment

MH thanks Prof. Masahide Sasaki, Prof. Akihisa Tomita, and Dr. Toyohiro Tsurumaru for valuable comments. He is partially supported by a MEXT Grant-in-Aid for Scientific Research (A) No. 23246071. He is also partially supported by the National Institute of Information and Communication Technology (NICT), Japan. The Center for Quantum Technologies is funded by the Singapore Ministry of Education and the National Research Foundation as part of the Research Centres of Excellence programme.

Appendix A. Chernoff inequality

In this section, we derive a lower bound of the lower percent point $Y^-(N, p, \alpha)$ with probability α by using Chernoff inequality. When the random variable X obeys the binomial distribution $\text{Bin}(N, p)$, Chernoff inequality

$$P_p\{X \leq Nq\} \leq \exp(-ND(q\|p)) \quad (\text{A.1})$$

holds with $q < p$, where the relative entropy $D(q\|p)$ is defined as $q \log \frac{q}{p} + (1-q) \log \frac{1-q}{1-p}$, where P_p is the distribution when the success probability with one trial is p .

Hence, letting q^- be the solution of the equation $D(q\|p) = -\frac{\log \alpha}{N}$ concerning q with $q < p$, we obtain

$$P_p\{X \leq Nq^-\} \leq \exp(-ND(q^-\|p)) = \alpha. \quad (\text{A.2})$$

That is, we obtain $Y^-(N, p, \alpha) \geq Nq^-$. Similarly, letting q^+ be the solution of the equation $D(q\|p) = -\frac{\log \alpha}{N}$ concerning q with $q > p$, we obtain $Y^+(N, p, \alpha) \leq Nq^+$.

Further, combining Pinsker inequality $D(q\|p) \geq 2(\log e)(p-q)^2$, we obtain

$$P_p\{X \leq Nq\} \leq \exp(-2(\log e)N(p-q)^2). \quad (\text{A.3})$$

Hence, solving the equation $2(\log e)(p-q)^2 = -\frac{\log \alpha}{N}$ concerning q , we obtain two solutions $\tilde{q}^- := p - \sqrt{\frac{-\log \alpha}{2(\log e)N}}$ and $\tilde{q}^+ := p + \sqrt{\frac{-\log \alpha}{2(\log e)N}}$. Then, we obtain $Y^-(N, p, \alpha) \geq N\tilde{q}^-$ and $Y^+(N, p, \alpha) \leq N\tilde{q}^+$.

Using the information geometry, we have a better evaluation than Pinsker inequality as follows. The relative entropy can be written with an integral form as follows[44].

$$\frac{D(q\|p)}{\log e} = \int_q^p \frac{t-p}{t(1-t)} dt. \quad (\text{A.4})$$

We consider only the case $p < 1/2$. When $q < p < 1/2$, we have

$$\frac{D(q\|p)}{\log e} \geq \frac{(p-q)^2}{2p(1-p)}. \quad (\text{A.5})$$

Hence, solving the equation $\frac{(p-q)^2}{2p(1-p)} = -\frac{\log \alpha}{N(\log e)}$ concerning q , we obtain the smaller solution $\bar{q}^- := p - \sqrt{\frac{-2(\log \alpha)p(1-p)}{(\log e)N}}$. Then, we obtain $Y^-(N, p, \alpha) \geq N\bar{q}^-$.

The treatment for $Y^+(N, p, \alpha)$ is a little complex. When $p < q \leq 1/2$, we have

$$\frac{D(q\|p)}{\log e} \geq \frac{(p-q)^2}{2q(1-q)}. \quad (\text{A.6})$$

Hence, solving the equation $\frac{(p-q)^2}{2q(1-q)} = -\frac{\log \alpha}{N(\log e)}$ concerning q , we obtain the larger solution $\bar{q}^+ := \frac{p - \log \alpha / (N \log e) + \sqrt{(-p^2 + p - \log \alpha / (2N \log e)) \cdot (-2 \log \alpha) / (N \log e)}}{1 - 2 \log \alpha / (N \log e)}$. Then, when $\bar{q}^+ \leq 1/2$, we obtain $Y^+(N, p, \alpha) \leq N\bar{q}^+$. Indeed, since \bar{q}^+ is complicated, we introduce a simpler upper bound. since $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$,

$$\begin{aligned} \bar{q}^+ \leq \hat{q}^+ &:= \frac{p - \log \alpha / (N \log e) + \sqrt{(-p^2 + p)(-2 \log \alpha) / (N \log e)} + \sqrt{(\log \alpha / (N \log e))^2}}{1 - 2 \log \alpha / (N \log e)} \\ &= \frac{p - 2 \log \alpha / (N \log e) + \sqrt{p(1-p)(-2 \log \alpha) / (N \log e)}}{1 - 2 \log \alpha / (N \log e)}. \end{aligned}$$

Then, when $\hat{q}^+ \leq 1/2$, we obtain $Y^+(N, p, \alpha) \leq N\hat{q}^+$.

Appendix B. One-sided interval estimation

Appendix B.1. One-sided interval estimation based of F distribution

We consider lower one-sided interval estimation with the confidential level $1 - \alpha$ when we observe the value k subject to the binomial distribution $Bin(N, p)$ with the trial N and probability p .

For this purpose, we fix an integer k and define the constants

$$n_1 := 2(N - k + 1), \quad n_2 := 2k, \quad f_1 := \frac{n_2(1-p)}{n_1 p}, \quad (\text{B.1})$$

it is known that the random variable $F(n_1, n_2)$ subject to F distribution with the freedom (n_1, n_2) satisfies

$$P\{F(n_1, n_2) > f_1\} = P_p\{X \geq k\} = \sum_{i=k}^N \binom{N}{i} p^i (1-p)^{N-i}. \quad (\text{B.2})$$

Our task is solving $P_p\{X \geq k\} = 1 - \alpha$ concerning p with $p < \frac{k}{N}$ for a given k . Define f_1^* to be the solution of $P\{F(n_1, n_2) > f_1\} = 1 - \alpha$ concerning f_1 . Then, the solution $p = \frac{n_2}{n_1 f_1^* + n_2}$ satisfies the equation $\frac{n_2(1-p)}{n_1 p} = f_1^*$. Thus, we obtain

$$P_{\frac{n_2}{n_1 f_1^* + n_2}}\{X \geq k\} = 1 - \alpha. \quad (\text{B.3})$$

That is, $\frac{n_2}{n_1 f_1^* + n_2}$ is the lower confidence limit $p^{(-)}(N, k, \alpha)$ of the lower one-sided interval estimation with the confidential level $1 - \alpha$ when we observe the value k .

Similarly, we fix an integer k and define the constants

$$m_1 = 2(k + 1), \quad m_2 = 2(N - k), \quad f_2 = \frac{m_1 p}{m_2 (1-p)}, \quad (\text{B.4})$$

it is known that the random variable $F(m_1, m_2)$ subject to F distribution with the freedom (m_1, m_2) satisfies

$$P\{F(m_1, m_2) > f_2\} = P_p\{X \geq k\} \quad (\text{B.5})$$

Our task is solving $P_p\{X \geq k\} = \alpha$ concerning p with $p < \frac{k}{N}$ for a given k . Define f_2^* to be the solution of $P(F(m_1, m_2) > f_2) = \alpha$ concerning f_2 . Then, the solution $p = \frac{m_1 f_2}{m_1 f_2 + m_2}$ satisfies the equation $\frac{m_1 p}{m_2 (1-p)} = f_2^*$. Thus, we obtain

$$P_{\frac{m_2}{m_1 f_2^* + m_2}}\{X \geq k\} \left(1 - \frac{m_2}{m_1 f_2^* + m_2}\right)^{N-i} = \alpha. \quad (\text{B.6})$$

That is, $\frac{m_2}{m_1 f_2^* + m_2}$ is the upper confidence limit $p^{(+)}(N, k, \alpha)$ of the upper one-sided interval estimation with the confidential level $1 - \alpha$ when we observe the value k .

Appendix B.2. Application of Chernoff inequality

Assume that we observe the value k subject to the binomial distribution $Bin(N, p)$ with the trial N and probability p .

For a fixed integer k , we have

$$P_p\left\{\frac{X}{N} \leq \frac{k}{N}\right\} \leq \exp(-ND(\frac{k}{N} \| p)) \quad (\text{B.7})$$

with $\frac{k}{N} < p$. Hence, letting p^- be the solution of the equation $D(\frac{k}{N}||p) = -\frac{\log \alpha}{N}$ concerning p with $\frac{k}{N} < p$, we obtain

$$P_{p^-}\{\frac{X}{N} \leq \frac{k}{N}\} \leq \exp(-ND(\frac{k}{N}||p^-)) = \alpha. \quad (\text{B.8})$$

Thus, $p^- \leq p^{(-)}(N, k, \alpha)$.

Similarly, letting q^+ be the solution of the equation $D(\frac{k}{N}||p) = -\frac{\log \alpha}{N}$ concerning p with $\frac{k}{N} > p$, we obtain $p^+ \geq p^{(+)}(N, k, \alpha)$.

Further, combining Pinsker inequality $D(q||p) \geq 2(\log e)(p - q)^2$, we obtain

$$P_p\{\frac{X}{N} \leq \frac{k}{N}\} \leq \exp(-2(\log e)N(p - \frac{k}{N})^2). \quad (\text{B.9})$$

Hence, solving the equation $2(\log e)(p - \frac{k}{N})^2 = -\frac{\log \alpha}{N}$ concerning p , we obtain two solutions $\tilde{p}^- := \frac{k}{N} - \sqrt{\frac{-\log \alpha}{2(\log e)N}}$ and $\tilde{p}^+ := \frac{k}{N} + \sqrt{\frac{-\log \alpha}{2(\log e)N}}$. Then, we obtain $\tilde{p}^- \leq p^{(-)}(N, k, \alpha)$ and $\tilde{p}^+ \geq p^{(+)}(N, k, \alpha)$.

Using the relation (A.4), we consider better bounds only for the case $\frac{k}{N} < 1/2$. Solving the equation $\frac{(p - \frac{k}{N})^2}{2\frac{k}{N}(1 - \frac{k}{N})} = -\frac{\log \alpha}{N(\log e)}$ concerning p with $p < \frac{k}{N} < 1/2$, we obtain the smaller solution $\bar{p}^- := \frac{k}{N} - \sqrt{\frac{-2(\log \alpha)\frac{k}{N}(1 - \frac{k}{N})}{(\log e)N}}$. Then, we obtain $p^{(-)}(N, k, \alpha) \geq \bar{p}^-$.

The treatment for $p^{(+)}(N, k, \alpha)$ is a little complex. When $\frac{k}{N} < p \leq 1/2$, we have

$$\frac{D(\frac{k}{N}||p)}{\log e} \geq \frac{(p - q)^2}{2q(1 - q)}. \quad (\text{B.10})$$

Hence, solving the equation $\frac{(p - \frac{k}{N})^2}{2p(1 - p)} = -\frac{\log \alpha}{N(\log e)}$ concerning p , we obtain the larger solution $\bar{p}^+ := \frac{k/N - \log \alpha/(N \log e) + \sqrt{(-(k/N)^2 + k/N - \log \alpha/(2N \log e))(-2 \log \alpha)/(N \log e)}}{1 - 2 \log \alpha/(N \log e)}$. Then, when $\bar{p}^+ \leq 1/2$, we obtain $p^{(+)}(N, k, \alpha) \leq \bar{p}^+$. Indeed, since \bar{p}^+ is complicated, we introduce a simpler upper bound:

$$\bar{p}^+ \leq \hat{p}^+ := \frac{k/N - 2 \log \alpha/(N \log e) + \sqrt{k/N(1 - k/N)(-2 \log \alpha)/(N \log e)}}{1 - 2 \log \alpha/(N \log e)}.$$

Then, when $\hat{p}^+ \leq 1/2$, we obtain $p^{(+)}(N, k, \alpha) \leq \hat{p}^+$.

Appendix C. Asymptotic case

Next, we discuss the asymptotic behavior of $Y^\pm(N, p, \alpha)$, $p^{(\pm)}(N, k, \alpha)$, and $X^\pm(N, k, \alpha)$. When $p < 1/2$, Hence, even if p depends on N and is written by p_N , if p_N converges to a non-zero constant p_a , using \tilde{p}^- and \tilde{p}^+ , we obtain $Y^-(N, p_N, \alpha) \geq Np_N - \sqrt{\frac{-2(\log \alpha)p_a(1 - p_a)N}{(\log e)}} + o(\sqrt{N})$ and $Y^+(N, p_N, \alpha) \leq Np_N + \sqrt{\frac{-2(\log \alpha)p_a(1 - p_a)N}{(\log e)}} + o(\sqrt{N})$. If p_N converges to zero, $Y^-(N, p_N, \alpha) = Np_N + o(\sqrt{N})$ and $Y^+(N, p_N, \alpha) = Np_N + o(\sqrt{N})$. In fact, if p_N converges to a non-zero constant p_a , using the central limit theorem, we obtain $Y^-(N, p_N, \alpha) = Np_N - \sqrt{p_a(1 - p_a)N}x + o(\sqrt{N})$ and $Y^+(N, p_N, \alpha) = Np_N + \sqrt{p_a(1 - p_a)N}x + o(\sqrt{N})$, where x is defined as $\alpha = \int_x^\infty \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} dx$.

Similarly, if $\frac{k_N}{N}$ converges to a non-zero constant p_a , we obtain $X^-(N, k_N, \alpha) = k_N - \sqrt{p_a(1-p_a)N}x + o(\sqrt{N})$ and $X^+(N, k_N, \alpha) = k_N + \sqrt{p_a(1-p_a)N}x + o(\sqrt{N})$. If p_N converges to zero, $X^-(N, k_N, \alpha) = k_N + o(\sqrt{N})$ and $X^+(N, k_N, \alpha) = k_N + o(\sqrt{N})$.

Appendix D. Calculation with the Gaussian case

In order to calculate the sacrifice bit length given in Subsection 12.1, we need $E[e^{\mu_i}]$, $E[\mu_i e^{\mu_i}]$, $E[\mu_i^2 e^{\mu_i}]$, and ω_2 . For this purpose, we calculate $e^{-\frac{1}{2\sigma^2}(x-(\mu-\sigma^2))^2}$ as follows.

$$\frac{de^{-\frac{1}{2\sigma^2}(x-(\mu-\sigma^2))^2}}{dx} = -\frac{1}{\sigma^2}(x-(\mu-\sigma^2))e^{-\frac{1}{2\sigma^2}(x-(\mu-\sigma^2))^2} \quad (D.1)$$

$$\begin{aligned} \frac{d^2e^{-\frac{1}{2\sigma^2}(x-(\mu-\sigma^2))^2}}{dx^2} &= \frac{1}{\sigma^4}(x^2e^{-\frac{1}{2\sigma^2}(x-(\mu-\sigma^2))^2} - 2(\mu-\sigma^2)xe^{-\frac{1}{2\sigma^2}(x-(\mu-\sigma^2))^2} \\ &\quad + ((\mu-\sigma^2)^2 - \sigma^2)e^{-\frac{1}{2\sigma^2}(x-(\mu-\sigma^2))^2}) \end{aligned} \quad (D.2)$$

Hence, $xe^{-\frac{1}{2\sigma^2}(x-(\mu-\sigma^2))^2}$, $x^2e^{-\frac{1}{2\sigma^2}(x-(\mu-\sigma^2))^2}$ can be written by using $e^{-\frac{1}{2\sigma^2}(x-(\mu-\sigma^2))^2}$ and its first and second derivatives as follows.

$$xe^{-\frac{1}{2\sigma^2}(x-(\mu-\sigma^2))^2} = -\sigma^2 \frac{de^{-\frac{1}{2\sigma^2}(x-(\mu-\sigma^2))^2}}{dx} + (\mu-\sigma^2)e^{-\frac{1}{2\sigma^2}(x-(\mu-\sigma^2))^2} \quad (D.3)$$

$$x^2e^{-\frac{1}{2\sigma^2}(x-(\mu-\sigma^2))^2} = \sigma^4 \frac{d^2e^{-\frac{1}{2\sigma^2}(x-(\mu-\sigma^2))^2}}{dx^2} + (2(\mu-\sigma^2)x - ((\mu-\sigma^2)^2 - \sigma^2))e^{-\frac{1}{2\sigma^2}(x-(\mu-\sigma^2))^2}. \quad (D.4)$$

We also prepare the following formula for $e^{-x}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$.

$$\begin{aligned} e^{-x}e^{-\frac{(x-\mu)^2}{2\sigma^2}} &= e^{-\frac{(x-\mu)^2}{2\sigma^2}-x} = e^{-\frac{1}{2\sigma^2}(x^2-2(\mu-\sigma^2)x+\mu^2)} \\ &= e^{-\frac{1}{2\sigma^2}(x-(\mu-\sigma^2))^2} e^{-\frac{\mu^2}{2\sigma^2} + \frac{(\mu-\sigma^2)^2}{2\sigma^2}} = e^{-\frac{1}{2\sigma^2}(x-(\mu-\sigma^2))^2} e^{\frac{(\sigma^2-2\mu)}{2}}. \end{aligned} \quad (D.5)$$

When X obeys the Gaussian distribution with the average μ and the variance σ^2 , using (D.3), (D.4), and (D.5), we can calculate the expectations of e^{-x} , xe^{-x} , and x^2e^{-x} as follows.

$$\begin{aligned} E[e^{-x}] &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-x}e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \frac{e^{\frac{(\sigma^2-2\mu)}{2}}}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}(x-(\mu-\sigma^2))^2} dx \\ &= e^{\frac{(\sigma^2-2\mu)}{2}} \end{aligned} \quad (D.6)$$

$$\begin{aligned} E[xe^{-x}] &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} xe^{-x}e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \frac{e^{\frac{(\sigma^2-2\mu)}{2}}}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} xe^{-\frac{1}{2\sigma^2}(x-(\mu-\sigma^2))^2} dx \\ &= e^{\frac{(\sigma^2-2\mu)}{2}}(\mu-\sigma^2) \end{aligned} \quad (D.7)$$

$$\begin{aligned} E[x^2e^{-x}] &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} x^2e^{-x}e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \frac{e^{\frac{(\sigma^2-2\mu)}{2}}}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} x^2e^{-\frac{1}{2\sigma^2}(x-(\mu-\sigma^2))^2} dx \\ &= \frac{e^{\frac{(\sigma^2-2\mu)}{2}}}{\sqrt{2\pi\sigma^2}} \left(\sigma^4 \int_{-\infty}^{\infty} \frac{d^2e^{-\frac{1}{2\sigma^2}(x-(\mu-\sigma^2))^2}}{dx^2} dx + 2(\mu-\sigma^2) \int_{-\infty}^{\infty} xe^{-\frac{1}{2\sigma^2}(x-(\mu-\sigma^2))^2} dx \right. \\ &\quad \left. - ((\mu-\sigma^2)^2 - \sigma^2) \int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}(x-(\mu-\sigma^2))^2} dx \right) \end{aligned}$$

$$= e^{\frac{(\sigma^2 - 2\mu)}{2}} ((\mu - \sigma^2)^2 + \sigma^2). \quad (\text{D.8})$$

Next, we calculate the real number ω_2 when μ_1 obeys the Gaussian distribution with the average μ and the variance σ^2 .

$$\begin{aligned} \omega_2 &:= \sum_{n=2}^{\infty} \frac{\mathbb{E}[e^{-\mu_1} \mu_1^n]}{n! \mathbb{E}[e^{-\mu_1} \mu_1^2]} = \frac{1}{\mathbb{E}[e^{-\mu_1} \mu_1^2]} \mathbb{E}[e^{-\mu_1} \sum_{n=2}^{\infty} \frac{1}{n!} \mu_1^n] \\ &= \frac{1}{\mathbb{E}[e^{-\mu_1} \mu_1^2]} \mathbb{E}[e^{-\mu_1} ((\sum_{n=0}^{\infty} \frac{1}{n!} \mu_1^n) - 1 - \mu_1)] = \frac{1}{\mathbb{E}[e^{-\mu_1} \mu_1^2]} \mathbb{E}[e^{-\mu_1} (e^{\mu_1} - 1 - \mu_1)] \\ &= \frac{1}{\mathbb{E}[e^{-\mu_1} \mu_1^2]} (1 - \mathbb{E}[e^{-\mu_1}] - \mathbb{E}[\mu_1 e^{-\mu_1}]) = \frac{e^{-\frac{(\sigma^2 - 2\mu)}{2}} - (\mu - \sigma^2) - 1}{(\mu - \sigma^2)^2 + \sigma^2}. \end{aligned} \quad (\text{D.9})$$

Appendix E. Relation with Eve's success probability

We consider the state $\rho_{AE} := \sum_m P(m) |m\rangle \langle m| \otimes \rho_{AE|m}$, where $\rho_{AE|m}$ is the composite state on $(\mathbb{C}^2)^{\otimes m} \otimes \mathcal{H}_E$. Now, we consider a function f from $\cup_m \{0, 1\}^m$ to $\{0, 1\}$. Then, we have the state $\rho_{f(A), E} = \sum_m P(m) \rho_{f(A)E|m}$ on $\mathbb{C}^2 \otimes \mathcal{H}_E$. Due to the monotonicity of the trace norm, the state $\rho_{f(A), E}$ satisfies

$$\|\rho_{f(A), E} - \rho_{f(A)} \otimes \rho_E\| \leq \|\rho_{A, E} - \rho_{\text{ideal}}\|. \quad (\text{E.1})$$

When $\rho_{f(A), E} = p_0 |0\rangle \langle 0| \otimes \rho_{0, E} + p_1 |1\rangle \langle 1| \otimes \rho_{1, E}$, due to the monotonicity of the trace norm, any two-valued POVM $\{T, I - T\}$ on \mathcal{H}_E satisfies

$$\begin{aligned} &\|\rho_{f(A), E} - \rho_{f(A)} \otimes \rho_E\| \\ &\geq p_0 (|\text{Tr } \rho_{0, E} T - \text{Tr}(p_0 \rho_{0, E} + p_1 \rho_{1, E}) T| + |\text{Tr } \rho_{0, E} (I - T) - \text{Tr}(p_0 \rho_{0, E} + p_1 \rho_{1, E}) (I - T)|) \\ &\quad + p_1 (|\text{Tr } \rho_{1, E} T - \text{Tr}(p_0 \rho_{0, E} + p_1 \rho_{1, E}) T| + |\text{Tr } \rho_{1, E} (I - T) - \text{Tr}(p_0 \rho_{0, E} + p_1 \rho_{1, E}) (I - T)|) \\ &= 4p_0 p_1 |\text{Tr } \rho_{0, E} T - \text{Tr } \rho_{1, E} T|. \end{aligned}$$

When T supports $f(A) = 0$ and $I - T$ supports $f(A) = 1$, the success probability is bounded by

$$\begin{aligned} &p_0 \text{Tr } \rho_{0, E} T + p_1 \text{Tr } \rho_{1, E} (I - T) \leq \max(p_0, p_1) (\text{Tr } \rho_{0, E} (I - T) + \text{Tr } \rho_{1, E} T) \\ &= \max(p_0, p_1) (1 + |\text{Tr } \rho_{0, E} T - \text{Tr } \rho_{1, E} T|) \leq \min(p_0, p_1) (1 + \frac{1}{4p_0 p_1} \|\rho_{f(A), E} - \rho_{f(A)} \otimes \rho_E\|). \end{aligned}$$

References

- [1] Bennett C H and Brassard G 1984 *Proc. IEEE Int. Conf. on Computers Systems and Signal Processing (Bangalore, India)* (New York: IEEE) pp 175–179
- [2] Shor P W and Preskill J 2000 *Phys. Rev. Lett.* **85** 441
- [3] Mayers D 2001 *Journal of the ACM* **48** 351
- [4] Watanabe S, Matsumoto R, and Uyematsu T 2006 *Int. J. Quant. Infor.* **4** 935
- [5] Hayashi M 2006 *Phys. Rev. A* **74** 022307
- [6] Gottesman D, Lo H-K, Lütkenhaus N, and Preskill J 2004 *Quant. Inf. Comput.* **5** 325 - 360
- [7] Hwang W-Y 2003 *Phys. Rev. Lett.* **91** 057901
- [8] Lo H-K, Ma X, and Chen K 2005 *Phys. Rev. Lett.* **94** 230504

- [9] Wang X-B 2005 *Phys. Rev. Lett.* **94** 230503
- [10] Hayashi M 2007 *New J. Phys.* **9** 284
- [11] Tsurumaru T, Soujaeff A, and Takeuchi S 2008 *Phys. Rev. A* **77** 022319
- [12] Curty M, Moroder T, Ma X, Lo H-K, and Lütkenhaus N 2009 *Phys. Rev. A* **79** 032335
- [13] Wang X-B, Yang L, Peng C-Z, and Pan J-W 2009 *New J. Phys.* **11** 075006
- [14] Wang X-B, Peng C-Z, Zhang J, Yang L, and Pan J-W 2008 *Phys. Rev. A* **77** 042311
- [15] Renner R 2005 *Security of Quantum Key Distribution* PhD thesis, Dipl. Phys. ETH, Switzerland; (eprint arXiv:quantph/0512258)
- [16] Hayashi M 2011 *IEEE Trans. Inform. Theory* **57** 3989
- [17] Hayashi M 2011 eprint arXiv:1010.1358
- [18] Hayashi M 2012 eprint arXiv:1202.0322
- [19] Wegman M N and Carter J L 1981 *J. Comput. System Sci.* **22** 265
- [20] Miyadera T 2006 *Phys. Rev. A* **73** 042317
- [21] Hayashi M 2009 *Phys. Rev. A* **79** 019901(E)
- [22] Koashi M 2009 *New J. Phys.* **11** 045018
- [23] Renes J M 2011 *Proc. R. Soc. A* **467** 1604
- [24] Tsurumaru T and Hayashi M 2011 eprint arXiv:1101.0064 (Accepted for publication for *IEEE Trans. Inform. Theory*.)
- [25] Strassen V 1962 Asymptotische Abschätzungen in Shannon's Informationstheorie In *Transactions of the Third Prague Conference on Information Theory etc*, Czechoslovak Academy of Sciences, Prague, pp. 689-723
- [26] Hayashi M 2008 *IEEE Trans. Inform. Theory* **54** 4619
- [27] Hayashi M 2009 *IEEE Trans. Inform. Theory* **55** 4947
- [28] Polyanskiy Y, Poor H V, and Verdú S 2010 *IEEE Trans. Inform. Theory* **56** 2307
- [29] Scarani V and Renner R 2008 *Phys. Rev. Lett.* **100** 200501
- [30] Sano Y, Matsumoto R, and Uyematsu T 2010 *J. Phys. A* **43** 495302
- [31] Tomamichel M, Lim C C W, Gisin N, and Renner R 2012 *Nat. Commun.* **3** 634
- [32] Hayashi M and Tsurumaru T 2012 *New J. Phys.* **14** 093014
- [33] Furrer F, Franz T, Berta M, Leverrier A, Scholz V B, Tomamichel M, and Werner R F 2012 *Phys. Rev. Lett.* **109** 100502
- [34] Sasaki M, Fujiwara M, Ishizuka H, Klaus W, Wakui K, Takeoka M, Tanaka A, Yoshino K, Nambu Y, Takahashi S, Tajima A, Tomita A, Domeki T, Hasegawa T, Sakai Y, Kobayashi H, Asai T, Shimizu K, Tokura T, Tsurumaru T, Matsui M, Honjo T, Tamaki K, Takesue H, Tokura Y, Dynes J F, Dixon A R, Sharpe A W, Yuan Z L, Shields A J, Uchikoga S, Legré M, Robyr S, Trinkler P, Monat L, Page J-B, Ribordy G, Poppe A, Allacher A, Maurhart O, Länger T, Peev M and Zeilinger A 2011 *Opt. Express* **19** 10387
- [35] Stucki D, Legré M, Buntschu F, Clausen B, Felber N, Gisin N, Henzen L, Junod P, Litzistorf G, Monbaron P, Monat L, Page J-B, Perroud D, Ribordy G, Rochas A, Robyr S, Tavares J, Thew R, Trinkler P, Ventura S, Voirol R, Walenta N and Zbinden H 2011 *New J. Phys.* **13** 123001
- [36] Renner R and König R 2005 Universally composable privacy amplification against quantum adversaries *TCC: Theory of Cryptography: 2nd Theory of Cryptography Conference*, Lecture Notes in Computer Science vol 3378 ed J Kilian (Berlin: Springer) pp 407-25
- [37] Ben-Or M, Horodecki M, Leung D W, Mayers D and Oppenheim J 2005 The universal composable security of quantum key distribution *Theory of Cryptography: 2nd Theory of Cryptography Conf., TCC 2005* (Lecture Notes in Computer Science vol 3378) ed J Kilian (Berlin: Springer) pp 386-406
- [38] Fung C-H F, Ma X and Chau H F 2010 *Phys. Rev. A* **81** 012318
- [39] Stinson D R 1992 Universal hashing and authentication codes, in J. Feigenbaum (Ed.): *Advances in Cryptology - CRYPTO '91*, LNCS 576, pp.62-73
- [40] Haran R *Chernoff Bounds for Binomial and Hypergeometric Distributions*, <http://www.hariharan-ramesh.com/ppts/chernoff.pdf>.

- [41] Akaike H 1973 Information theory and an extension of the maximum likelihood principle, In B. N. Petrov and F. Csaki (Eds.), *Second international symposium on information theory* (pp. 267-281). Budapest: Akademiai Kiado.
- [42] Takeuchi K 1976 Distribution of information statistics and a criterion of model fitting *Suri-Kagaku* (Mathematical Sciences) **153** 12–18 [In Japanese]
- [43] Rissanen J 1978 *Automatica* **14** 465–471
- [44] Amari S and Nagaoka H 2000 *Methods of Information Geometry*, (AMS & Oxford University Press)